LIBRARY

FEB 24 1925

Vol. VIII, No. 1. *PSYCHOLOGICAL REVIEW PUBLICATIONS* FEBRUARY, 1925

Journal of Experimental Psychology

EDITED BY

JOHN B. WATSON, NEW YORK CITY
HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*Index*) AND
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)

ADVISORY EDITORS

H. A. CARR, UNIVERSITY OF CHICAGO; RAYMOND DODGE, WESLEYAN UNIVERSITY; KNIGHT DUNLAP, THE JOHNS HOPKINS UNIVERSITY; H. S. LANGFELD, PRINCETON UNIVERSITY; E. L. THORNDIKE, COLUMBIA UNIVERSITY.

CONTENTS

The Relation of Tongue Movements to Internal Speech: AGNES M. THORSON 1

Fatigue Tests and Incentives: HELEN FRANCIS WHITING AND HORACE BIDWELL ENGLISH 33

Relative Performance of College Students as Conditioned by Time of Day and Day of Week: DONALD A. LAIRD 50

Reactions to Visual Stimuli in Affective Settings: F. L. WELLS 64

PUBLISHED BI-MONTHLY BY
PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STREETS, LANCASTER, PA., AND
PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

Psychological Review Publications

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, 244 MADISON AVE., NEW YORK (*J. of Exp. Psych.*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*Index*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)

WITH THE CO-OPERATION OF

MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL BULLETIN

containing abstracts, critical reviews, of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in May, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

Philosophical Monographs: a series of treatises more philosophical in character.
Library of Genetic Science and Philosophy: a series of bound volumes.

ANNUAL SUBSCRIPTION RATES

Review: $5.00 (Foreign, $5.25). Review and Bulletin: $9.50 (Foreign, $10.00)
Journal: $5.00 (Foreign, $5.25). Review and Journal: $9.00 (Foreign, $9.50)
Bulletin: $5.50 (Foreign, $5.75). Journal and Bulletin: $9.50 (Foreign, $10.00)
Any one of above with Index: $1.50 additional.
Review, Bulletin, and Journal: $14.00 (Foreign, $14.75).
Review, Bulletin, Journal, and Index: $15.00 (Foreign, $15.75).
Current Numbers: Review or Journal, $1.00; Bulletin, 60c; Index, $2.00
Psychological Monographs: $6.00 per volume (Foreign, $6.30).
Current Issues: prices vary according to size.

Subscriptions, orders, and business communications may be sent direct to the

PRINCETON, N. J. OFFICE

PSYCHOLOGICAL REVIEW COMPANY

FOREIGN AGENTS: G. E. STECHERT & CO., London (2 Star Yard, Cary St., W.C.) Paris (16, rue de Condé)

Journal of
Experimental Psychology

VOL. VIII., No. 1 February, 1925

THE RELATION OF TONGUE MOVEMENTS TO INTERNAL SPEECH

BY AGNES M. THORSON

University of Minnesota

Since the discussion by Bain in 1879, the problem of the actual relation of 'internal speech' to verbal mechanisms has been of interest to psychologists, and has led to rather extensive discussion. Earlier writers are difficult to interpret as to their point of view concerning the function of overt movements. Bain ('79) made the statement that 'thinking is restrained speaking or acting.' He also says, "When we recall the impressions of a word or sentence, if we do not speak it out, we feel the twitter of the organs (vocal) . . . ." In the same year Ribot ('79) advanced a similar theory, "When man reads silently each visual perception is accompanied by suppressed articulation." He is of the opinion that all the psychical processes terminate in movement of some sort or other, but he does not particularly stress internal speech nor insist that it is necessary for all thinking. Most of the older theories of inner speech are characterized by the absence of any attempt to relate it to motor phenomena. Those of Egger ('81) and Ballet ('86) are the most conspicuous in this group. To these writers inner speech seems to be something intangible, something psychic, which they cannot explain, so they attempt to settle the controversy by calling that something inner speech. They do not wish to imply that inner speech is thinking, but that it is merely the vehicle of thought, or the outer garment in which thought makes its appearance. Inner speech does not enter in until a thought

is almost completely formed; therefore in thinking of something new and difficult the lapse of time between the genesis of thought and inner speech is long, while in thinking of something familiar thought and inner speech are almost simultaneous. Both Ballet and Egger rely mainly upon their own introspection in drawing their conclusions.

Stricker ('80) came to the conclusion that we could not have an idea of the sound 'B' without feeling some muscular movement. He also relied upon the introspective method, but did not restrict it to his own introspections. His method was to ask 100 persons whether when thinking or reciting silently a poem, they spoke to themselves. All answered 'yes.' Paulhan ('86) objected to Stricker's conclusion, since he maintained that he could have an image of a vowel while pronouncing aloud another vowel. He has no objection to considering thought an inner language, but he does not believe this language can be reduced to words or images of words. It is not clear what he does mean by 'inner language.'

These writers have pointed out the possible significance of actual vocal movements in thought, or in so-called internal speech, but have not stated very clearly the extent to which they believed them to play an essential part in internal speech.

Watson ('14) was the first to maintain the identity of thought and language activity. He states that no one has successfully connected thought with the rest of the process of integration, that investigators have for the most part assumed that, "it is something, no one knows quite what, that can go on in the absence apparently of all muscular activity." He states further, "we will suppose that future analysis will enable us ultimately to show that every word, syllable, and letter, whether spoken or thought, produces a characteristic form of response. . . ."

Throughout this discussion Watson uses thought as synonymous with the internal speech or imagery of earlier writers. This fact led to some confusion in the British Symposium ('11), when thinking was interpreted, particularly by Smith and Pear, as meaning the adaptive reactions of

the organism, whether involving verbal formulations or not. Watson restricts his identification of thinking and language mechanisms to verbal or symbolic formulation of ideas, rather than to mechanisms of reasoning or general adaptive reactions. To this extent the criticisms urged against him in the Symposium were aside from the point.[1]

Since the rôle of motor organs in internal speech is a problem which can be attacked experimentally, we must appeal to the actual analysis of movements to settle it. Certain experimental work has been attempted.

As early as 1895 Hansen and Lehman found that when a person concentrated greatly upon a word or number an unintentional whispering nearly always occurred, which could actually be heard by an observer. In other words, articulation accompanied all concentrated thought. These same investigators in another experiment on thought transference ('96) found some indication of the fact that the thought of one person could be interpreted by another by means of observable articulatory movements made during the thinking.

Curtis ('99) worked on the movements of the larynx. He placed a tambour on the larynx, while the subject recited or whispered familiar verses, or read to himself from a book. Out of 40 curves on 20 subjects, 18 showed clearly marked automatic movements of the larynx in some or all of the tracings, 11 showed smaller but still distinguishable variations, the remaining 11 showed hardly any difference between 'normals' and other curves. Out of 20 subjects 15 showed movements, and 5 none at all, not even in whispering.

Courten ('02) placed a rubber bulb on the tongue, connected with a recording tambour. Otherwise his method was similar to that of Curtis. He found that the amount of movement differed with the individual and also with the

[1] Owing to this confusion in terminology I shall use the term *internal speech* for the as yet obscure processes which are involved in the verbal formulation of thought, without prejudicing the case as to whether this internal speech is to be regarded as a succession of kinesthetic images or of actual overt movements. I shall use the term *implicit speech* for the actual slight muscular contractions which have been considered by Watson and others to constitute internal speech; and the term *overt speech* for voluntarily articulated language.

matter read or thought about, and the degree of attention. No record was taken which did not show some movement.

Wyczoikowski ('13) placed the tongue in a flattened glass cup connected with a recording tambour, while the subject thought of a familiar verse, the words 'experimental psychology' or listened to spoken words. Her conclusion was that every thought produced some movement of the tongue.

Pintner ('13) tested the degree to which the process of articulation could be eliminated in silent reading without lessening the degree of comprehension and rate of reading. His method was to have the subject read over certain selected passages of prose, while inhibiting articulatory movements. The inhibitory process was repeating la, la, la, or reciting numerals. His conclusions are:

1. Articulation during the reading process is a habit not necessary.

2. Practice in reading without articulation can make such reading as good as the ordinary.

3. Practice in reading without articulation tends to aid ordinary reading.

Reed ('16) experimented upon tongue movements. He placed in the subject's mouth a sensitive rubber bag, attached to a wooden framework placed between the teeth. This was connected with a tambour by means of a glass tube. The tambour was connected with a lever, which recorded movements on a kymograph drum. His stimuli were: silent reading, whispering, writing abstracts of newspaper clippings, mental multiplication, and the like. His results show in several cases definite tongue movements during silent reading comparable to those made during the whispering of the same material. He concludes that thinking is definitely a process of tongue movements for some subjects, but that this is an individual trait rather than a universal one.

Clark ('22) secured records of both laryngeal and tongue movements simultaneously, placing a tambour on the larynx and a rubber bulb on the tongue, while the subject was given a problem to solve. The results on tongue movements show the occasional presence of inner speech in each reaction; and

laryngeal movements were occasionally recorded when the tongue was especially active. She concludes that while inner speech plays an important part in thought, the greater part of thinking is carried on without it.

Most of these studies are open to three main objections:

1. *The type of apparatus used;* in practically all of the former experiments the rubber bulb type of apparatus was used to record tongue movements. Such apparatus can not possibly record tongue movements only, for any slight change in air pressure due to respiration, swallowing movements, and the like, would cause deflation of the rubber bulb, even though they were not great enough to bring about any movement of the tongue. This constitutes a constant error which is present during thinking and whispering alike. Its influence is possibly of sufficient magnitude to give a false appearance of similarity to curves of tongue or laryngeal movement.

2. *Interpretation of results:* It has been the common practice for investigators to interpret any movements of the tongue whatever as an indication of internal speech. The quantitative study made in this experiment (see results) shows this to be a fallacy.

3. *Suggestibility:* In all mechanical recording of movement the attachment of the apparatus, and the resultant discomfort tend to direct the subject's attention to the tongue and thus may tend to suggest movements. Subsequent taking of introspections may have a similar effect unless carefully controlled. In this respect Reed's technique seems especially open to criticism. As part of his experimental procedure, the subject was questioned, after each situation, as to whether or not he used inner speech during that situation. These introspective results were summarized and compared with the actual graphic records for the corresponding situations. This comparison reveals the following: 60 per cent. of the time there was absolute agreement between the subject's awareness of inner speech and the indication of it on the graphic records; 31 per cent. of the time the subject was not quite sure whether he used implicit movements or not, nor did the graphic records indicate it clearly; and only 9 per cent.

of the time was there disagreement between the introspective and graphic results. This very close agreement (91 per cent.) between the subject's awareness of the presence or absence of implicit movements and the objective results, seems to indicate that the functioning of the tongue as a mechanism for thinking is not a normal one, but an artificial one forced for the purpose of the experiment. Furthermore suggestible subjects, by way of this technique, would have their attention attracted to the tongue more than normally, and increased movement would follow.

In a more recent development of the theory, sketched by Watson in the British Symposium, a further problem was brought forth. That is, the possibility of the implicit activity shifting from one language mechanism to another. If the results of our experiments are to be valid, this shift must be controlled. However in the experiments cited above, little has been done to control this; and where it has been controlled to some extent, as in Pintner's work, the results have not been sufficiently analyzed.

Consequently this earlier experimental work must be regarded as inconclusive, and the problem is still unsettled as to whether internal and implicit speech ought to be looked upon as identical; or whether internal speech ought to be regarded as representing some form of wholly central neural activity not dependent upon implicit speech.

The experiments reported in this study were designed to test the following:

1. Do movements of the tongue occur constantly or with any significant frequency during internal speech?

2. Are the movements which occur identical in their form to those of the corresponding overt speech?

3. If so, are they an essential element of internal speech or only an incidental consequence of it?

4. If they do not occur regularly, what is the significance of the occasional movements which have been reported?

Apparatus

The apparatus used for recording tongue movements in this experiment was devised by Dr. K. S. Lashley, and is a modification of Sommer's 'Movement analyzer.' [1] This apparatus is shown in diagram form in Fig. 1. *A*, *B*, and *C* are upright supports in which the recording levers are mounted. A small metal suction cup *i* is attached to the tongue. This cup is attached to a lever system at *g* by means of a fine wire *h*, which passes between the teeth. Tongue movements in a vertical plane are communicated to the lower writing point *f* by means of the bell cranks *a* and *b*. Horizontal movements of the tongue are communicated through levers *a*, *c*, and *d*

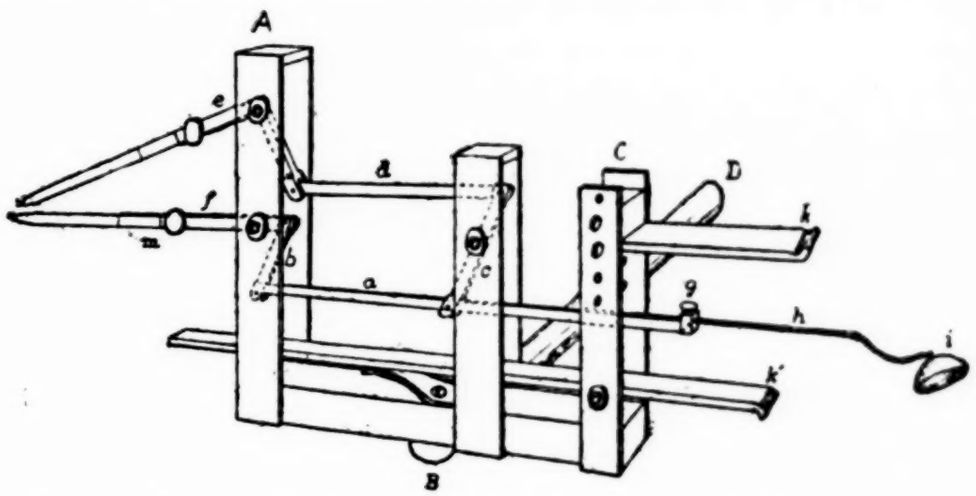


Fig. 1. Diagram of instrument used to record movements of the tongue, × 1/2. *ABC*, rigid frame supported in adjustable clamp by the rod, *BD*. *i*, metal suction cup for attachment to tongue. *a*, aluminum bar, attached to suction cup by the aluminum wire, *h*, adjustable by the set screw, *g*, and transmitting movements through the links and bell-cranks, *b*, *c*, *d*, and *e*, to the writing points. The writing points are adjustable horizontally by thumb screws, *m*, acting against light bronze springs.

to the bell crank *e* which carries the upper writing point. The writing points *e* and *f* are arranged so as to record on smoked paper.[2]

The upper recording point magnifies the movements approximately 3.5 times, and the lower recording point

[1] I wish to acknowledge my indebtedness to Dr. K. S. Lashley for the suggestion of the problem, construction of the apparatus, and for his direction and supervision of the work throughout.

[2] Extension arms *k* and *k'* were designed as rests for the teeth but were not used in this experiment.

magnifies approximately 4.5 times. Therefore an actual tongue movement of one fourth or one fifth of a millimeter can easily be recorded and measured on the curve.

The particular advantage of this apparatus is the fact that it is not sensitive to the influence of respiration and of swallowing movements except when these directly affect the tongue. It also simplifies the records by separating the vertical and horizontal movements into two distinct curves. This permits of more ready identification.

A long-paper-kymograph which advanced the smoked paper nine mm. per second was used. It provided continuous records for about 15 minutes. The time was recorded in seconds with a Jacquet timer. A cardboard screen was placed in front of the subject's eyes, in order that any factor of suggestion from watching the movements of the levers and the tracings made on the drum might be eliminated. To minimize head movements, the subject's head was strapped into a head rest, which was free from any chin or laryngeal obstruction. A telegraph key connected with a signal magnet provided the subject with a means of indicating on the kymograph record, the beginning of each new repetition of a task. During the major part of the experiment tracings of thoracic respiration were taken by means of a Sommer pneumograph.

Method

The following experimental procedure was used. The subject was seated on a revolving stool, his head strapped securely in the head rest, and the suction cup attached to his tongue. Then adjustments were made so that all levers tracing upon the drum were parallel, the tips of all forming a vertical straight line upon the drum. Thus the curved tracings could be referred to a constant base line and so compared for duration. With only the fine wire serving as a connection between the suction cup on the tongue and the rest of the apparatus, it was not necessary for the subject to keep his mouth open more than a few millimeters. Of course, his lower jaw must be relaxed and his teeth apart, in order

not to interfere with correct tracings of movements, but aside from this the situation was a close approximation to normal and not necessarily uncomfortable. Even whispering was possible without a great deal of interference by the apparatus, except when forming the dental sounds.

Complete instructions were given to the subject verbally, before the kymograph was set running, and immediately preceding each new task. The subject was instructed to press a telegraph key at each repetition of the task.

Five different groups of tasks were used, one at each sitting of the subject. Each group was selected for the purpose of solving a definite problem.

The problems presented themselves as follows:

1. It was found desirable to determine the similarity between movements in the repetition of the same internal speech, and to compare these with the corresponding movements in overt speech. For the identification and analysis of similar movements a simple task was required. Groups I. and II. were selected for this purpose. In group I. each task was thought slowly and repeated several times (except parts 3 and 4). In group II. each task was whispered and then thought (except part 3), repeating the process several times.

2. It was also desirable to make the determination of movements with more difficult material. Groups III. and IV. were chosen for this purpose. In group III. a very difficult passage in Polish was whispered and thought alternately. In group IV. a short newspaper clipping was written rapidly from dictation, read silently, and read in a whisper.

3. Lastly, it was of interest to determine the effect of distraction and emotional disturbance on movements. Also it was desirable to eliminate to some extent the possibility of shift of internal speech to laryngeal, manual, and respiratory mechanisms. This led to the introduction of singing a constant tone, tapping with the fingers, and the taking of respiration records. As a basis for this study the tasks of group V. were used. In group V. the same tasks were repeated with and without distraction.

As the problem developed, the amount of movement irrespective of its pattern seemed significant. So controls were introduced, in which the subjects were instructed to inhibit internal speech, close the eyes, and try to sleep. These situations are designated as 'blanks' throughout.

A detailed statement of the separate tasks (material to be repeated in internal or overt speech) in each group follows: the nonsense material was presented visually, to be read and pronounced in internal speech.

Group I:

1. 'Experimental psychology.'
2. 'Shave a cedar shingle thin.'
3. Mental multiplication—'25 x 23.'
4. Mental multiplication—'69 x 72.'
5. 'Duka rima rinka ro.'
6. 'Bonim bokie salib so.'
7. 'Experimental psychology.'
8. Blank.

Group II:

1. 'Individual differences.'
2. 'If she sells sea shells by the seashore, why can't he sell sea shells by the seashore.'
3. Mental addition of 12 three place numbers.
4. Blank.
5. 'Raemenpuimembimer.'
6. 'Sam lip rascal nes lo tod.'
7. 'Entac riplam in sel rich.'
8. 'Jack and Jill went up the hill.'
9. 'Experimental psychology.'

Group III:

1. "Palacy! Kasciuszko i pulaski walczyli zu walnosc Polski i innych narodow! dd'zmy wich slady! Hej na boj z wrogum odwucznym Polski i walmas ci!"

Group IV:

1. A short newspaper clipping.

Group V:

1. Thinking 'experimental psychology' while constantly singing 'ah,' and tapping the table with the fingers.

2. Thinking 'experimental psychology' without any distraction.

3. Reciting silently the multiplication table of 7's or 9's from 1 to 14 while singing 'ah,' and tapping.

4. Reciting the multiplication table of 7's or 9's from 1 to 14 without any distraction.

5. Reciting silently the entire alphabet as rapidly as possible.

6. Multiplying mentally a two place number by a two place number, while the experimenter rasps irregularly with a coarse file upon a piece of metal clamped to a table (suggested by Landis ('24)).

7. Blank.

Controls.—As the experiment progressed, there was some indication that tongue movements may be produced as a result of several different sorts of reactions, such as breathing, swallowing, excessive salivation, tongue tremor, and increased nervous tension. The first of these is common to all situations, but since all conclusions are based on relative amounts of movement, this becomes negligible. The swallowing was taken care of by careful watching of the subject, so that whenever this occurred, with its characteristic mark on the curve, the record was discounted in computing results. When excessive salivation occurred, and this was present in only a very few subjects, the subject was permitted to swallow and the situation was repeated, the first results being thrown out.

The amount of movement due to tongue tremor was measured by control tests referred to already as 'blanks.' Despite the fact that a few subjects reported having difficulty in keeping their 'thoughts blank,' there is a high degree of uniformity of the results obtained on the different subjects, and the average amount of movement during these various blank situations is lower than that of the average in any other situation.

The amount of movement due to variations in nervous tension is measured quantitatively by repeating the same and similar stimuli under conditions with and without increased nervous tension.[1]

Subjects: Eighteen subjects were used in this experiment, 14 of whom were faculty members or graduate students in the department of psychology. Three undergraduate students and one ten-year-old boy were also used. The boy could be secured for one sitting only, but his results are very similar to the rest obtained, so they are included here.

The Relation of Pattern of Movement to Internal Speech

The primary interest of the problem lies in the possible identity of implicit movements and inner speech, or 'thought' as used by Watson ('11). If the theory of implicit speech is true, and if the tongue plays an important part in it, we should find some similarity between the movements of the tongue which appear in overt speech and in implicit speech during the repetition of simple words or phrases.

To test this, first, a comparison was made of actual whispering movements with the movements occurring during a repetition of the whispered phrases in internal speech. The results of group II. were analyzed to determine the number of peaks, depressions, and changes in directions of movement for each curve obtained during overt speech (whispering) and the corresponding ones obtained during internal speech. The material of group II. consists mainly of short phrases (see page 10), making a word for word analysis and comparison fairly simple. 131 pairs of curves (one for whispering and one for internal speech) for these tasks were compared. They were considered similar when any possible similarity between them could be noted; such as having the same number of major changes in direction, and

[1] The term 'nervous tension' is used, for lack of a better expression, to characterize the conditions of increased excitability which seem to appear after innervation of an organ, during emotional disturbance, or excessive effort. Physiologically the condition is probably one of facilitation of final common paths through irradiation. (Adrian '23, Head '23, Myers '23, and Pieron '23.)

the same number of major peaks and depressions, ignoring all the lesser peaks. Fig. 2 shows the most similar of the records obtained. The results of this analysis are shown in Table I. Of ten subjects, only three show any similarity at all between pattern of tongue movement in internal speech and in whispering, and these three show similarity in only a small per cent. of cases. The other 7 subjects show in 100

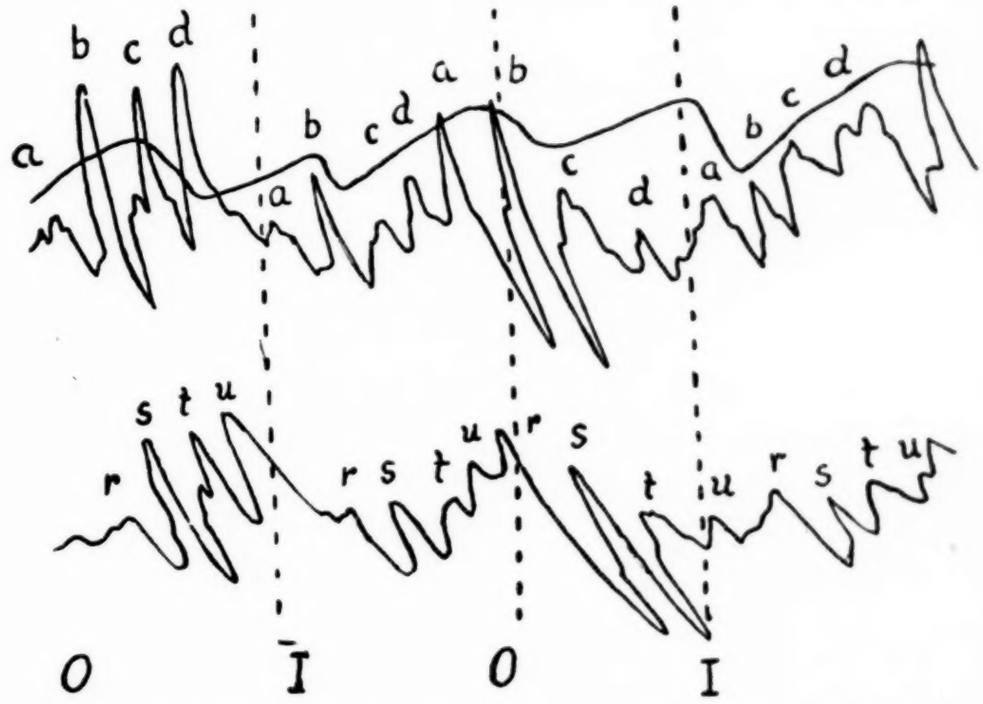


FIG. 2. Tracings of curves from subject *B*, alternate whispering and silent thinking of the phrase, "Jack and Jill went up the hill." These are among the most nearly alike of the curves included as 'similar' in Table I. Upper line, respiration; middle line, horizontal movements; lower line, vertical movements. The lower case letters indicate corresponding points on the curves; *O*, overt whispering; *I*, internal speech.

per cent. of cases no relationship whatsoever. Averaging the results of all the subjects together, we find that 4.4 per cent. of the total cases show some similarities, as against 95.6 per cent. which show no similarities at all between pattern of movement in thinking as compared with whispering.

But possibly the implicit processes are so abbreviated, short-circuited and economized, to use Watson's terminology, that they would be unrecognizable when compared with the explicit processes. For this reason the tongue movement

TABLE I

COMPARISON OF PATTERN OF TONGUE MOVEMENTS DURING WHISPERING WITH PATTERN OF MOVEMENTS DURING INTERNAL SPEECH

Subject	Similar		Dissimilar	
	No. of Pairs of Curves	%	No. of Pairs of Curves	%
K	2	28.6	5	71.5
D	0	0.0	14	100.0
H	0	0.0	14	100.0
M	0	0.0	12	100.0
L	0	0.0	14	100.0
A	2	15.4	11	84.6
G	0	0.0	16	100.0
N	0	0.0	15	100.0
D	0	0.0	14	100.0
B	2	16.7	10	83.3
Totals	6	4.4	125	95.6

curves on the successive repetitions in internal speech of the same words and phrases were compared. The results grouped under three headings, identical, similar, and dissimilar are shown in Table II. Fig. 3 shows the types of curves listed

TABLE II

COMPARISON OF TONGUE MOVEMENTS DURING REPETITION OF THE SAME TASKS IN INTERNAL SPEECH

Subject	Identical		Similar		Dissimilar	
	No.	%	No. of Pairs of Curves	%	No. of Pairs of Curves	%
K	0	0.0	2	28.6	5	71.4
D	0	0.0	1	7.1	13	92.0
B	0	0.0	4	33.3	8	66.6
H	0	0.0	0	0.0	14	100.0
M	0	0.0	2	16.7	10	83.3
D	0	0.0	0	0.0	14	100.0
L	0	0.0	0	0.0	14	100.0
A	0	0.0	2	15.4	11	84.7
G	0	0.0	2	12.5	14	87.5
N	0	0.0	4	26.6	11	73.4
Totals	0	0.0	17	13.0	114	87.0

as similar. When repeating the same tasks, not a single internal speech curve is identical in pattern with any other internal speech curve. Of the total cases for all subjects,

13 per cent. show some similarity, ranging by subjects from 0.0 to 33.3 per cent. similarity, whereas 87 per cent. show no possible pattern relationship. Similarity reaches a considerable proportion only in the case of subjects 'K,' 'B,' and 'N.' Subjects 'K' and 'B' have already shown some indications of similarity between patterns of movements in internal speech and in whispering, so it is not very surprising to find that their internal speech curves occasionally correspond to each other.

These results seem to prove that movements of the tongue are not an essential element in internal speech. For if they were, we should expect to find records of movements showing

FIG. 3. Tracings of successive curves of subject N repeating 'Individual differences' in internal speech. The specimens are the most nearly identical of those called similar in Table II. The letters mark corresponding points of the curves. Upper line horizontal, lower vertical movements.

constant patterns during successive repetitions of the same task. Even if internal speech is an abbreviated form of overt speech, there is no reason to believe that a different abbreviation would be used at each repetition of the same task during the sitting.

To make sure that the apparatus was capable of recording the movements, had they occurred, a similar study was made of the records for the successive repetitions in whispering of the same words and phrases. These results are shown in Table III. Typical curves for whispering are shown in Fig. 4.

Out of the total cases for all subjects, 25.2 per cent. are identical, 64.2 per cent. similar, and only in 10.6 per cent. of the cases is there no pattern relationship. This makes a total of 89 per cent. clearly showing pattern relationship.

The 10.6 per cent. not showing pattern relationship can be explained by the fact that the subjects were not always equally conscientious about following instructions, and would sometimes slur over the words. This is sufficient evidence to enable one to say with certainty that the apparatus can record such movements when they occur.

TABLE III

COMPARISON OF TONGUE MOVEMENTS DURING REPEATED WHISPERING OF THE SAME TASKS

Subject	Identical		Similar		Dissimilar	
	No.	%	No. of Pairs of Curves	%	No. of Pairs of Curves	%
K.....	2	28.6	5	71.4	0	0.0
D.....	2	14.3	10	71.4	2	14.3
B.....	4	33.3	8	66.6	0	0.0
H.....	6	42.8	7	50.0	1	7.2
M.....	3	25.0	9	75.0	0	0.0
D.....	0	0.0	14	100.0	0	0.0
L.....	0	0.0	13	93.0	1	7.0
A.....	2	15.4	5	38.5	6	46.2
G.....	7	43.8	8	50.0	1	6.2
N.....	7	46.7	5	33.3	3	20.0
Totals..	33	25.2	84	64.2	14	10.6

It has been suggested (Watson '19) that a shift can take place in the internal speech mechanism which is functional at any one instant. Therefore, since the tongue is encumbered with apparatus the process of internal speech might slip from this motor region to another. To control this possibility the following experiments were undertaken. The subject was asked to think 'experimental psychology' repeatedly, while singing 'ah' and tapping on the table with the finger tips (parts 1 and 3 of group V. situations). This must rule out, to a great extent, the larynx and the fingers as mechanisms for implicit language. Of course it does not exhaust the range of motor organs which might conceivably function in implicit speech, but it does rule out the more probable ones (the chief gestural mechanisms). If these organs had been functioning, their elimination should throw the activity back to the tongue, and we should expect to

find that the tongue movements in internal speech, during this situation at least, bear some resemblance to movements in overt speech or a greater resemblance to each other. The curves resulting from this experiment show no closer relationship to whispering curves than do those obtained when these controls are not operating.

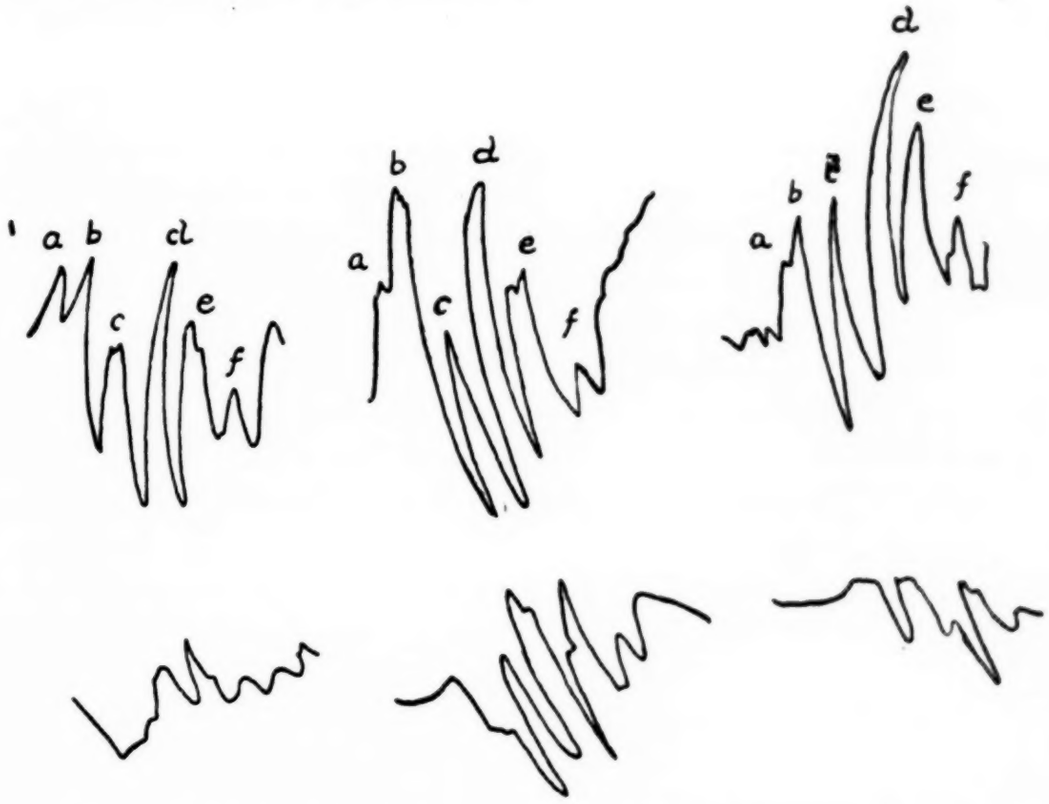


FIG. 4. Tracings of successive curves of subject N for overt whispering of 'individual differences' as an illustration of curves judged identical in Table III. The letters indicate corresponding points in the curve. Upper line horizontal, lower vertical movements.

If it is difficult to accept the statement that singing 'ah' rules out the larynx as a mechanism for internal speech, perhaps further evidence will help to substantiate the conclusion drawn. The larynx is capable of movement as a whole, and also the relation of its parts can be changed to bring about a change in tension of the vocal cords. But the only effects of the movements in overt speech are changes in the pitch and timbre of the voice. Therefore, if the larynx is functional during internal speech, any significant muscular contractions in it must be accompanied by changes in the

pitch of the voice. Any activity should therefore be recognized by a change in the pitch while singing 'ah,' but no such changes in pitch were noted during the experiment.

Furthermore, should the muscles of the thorax and diaphragm serve as a basis for implicit language, the respiration records must show some significant movements other than those produced by normal breathing. The respiration records obtained however showed no significant deviation from normal.

All of the records show that there is no correlation between the pattern of movements of the tongue in successive repetitions of the same words or phrases; and that this can not be interpreted by the assumption that internal speech is carried on by manual, laryngeal, or respiratory mechanisms.

The possibility still remains however, that movements are inhibited by attention to the apparatus. To direct the attention elsewhere, more difficult problems were introduced, such as writing from dictation, reading Polish, and solving problems under severe distractions. The results of these experiments show movements in varying amounts, but not even here is there any indication of pattern relationship.

The Relation of Amount of Movement to Internal Speech

Despite the lack of qualitative similarity, the records show continuous irregular movements of the tongue, varying in amount in different situations. What can be the significance of these increases in movement? If the increases were directly related to internal speech, they should show greater qualitative similarity under the conditions which produce them. Since they do not show this, they must be due to something else.

Possibly the increase is due to something in the way of differences in nervous tension. It was hoped that some light could be thrown on this problem by a study of the after effects of innervation of the tongue, by the introduction of several distractions which required an effort to carry on the problem, or by the introduction of emotional disturbances.

The following method was used in analyzing the data for this study. The actual amount of movement for each repetition of each task was measured in centimeters with a chartometer.[1]

Since the duration of the tasks and of internal speech varies, it becomes necessary to express the amount of movement as a ratio of extent to duration. To obtain these ratios the total amount of movement for each repetition of a task was divided by the straight line distance on the record between the beginning and end of the task. The averages of these ratios at each sitting for each subject for each task appear in the Appendix in Tables IX. to XIII. inclusive.[2]

Turning to a comparison of the amount of tongue movement during internal speech following whispering (situations of group II.) with the amount of movement during internal speech not preceded by whispering (situations of group I.), the averages for about five hundred ratios of both upper and lower curves show:

Averages for Group II	1.8 and 1.7
Averages for Group I	1.2 and 1.2
Difference	.6 and .5

Table IV. *A* [3] shows that 83.4 per cent. and 88.4 per cent. of the ratios of group II. are above the corresponding averages of group I. Table IV. *B* shows that 88 per cent. and 96 per cent. of the cases of group I. are below the averages of group II. This indicates that there is a very significant difference

[1] The wheel of this chartometer had been cut to about one third of normal size to facilitate the accurate measurement of very acute angles and minute peaks and depressions. The instrument was then recalibrated.

[2] The ratios in Table XIII. are the only exception to this system of recording results. Here the same tasks were repeated in the same ten subjects on a number of different days: five of them underwent the same experiment on six different days, and the figures appearing for them in the table are averages of six performances; the other five went through the experiment only twice, and the figures appearing for them in the table are averages of two performances. This was done as a check to determine the constancy of response to the same situations, from day to day. Since the subjects showed relatively the same results from one day to the next (although absolutely there were some differences) it was not considered necessary to make further repetitions.

[3] Percentages were used in the determination of the significance of the difference between two averages rather than the P.E.$_D$. Since these data are made up of averages of averages, the computation of P.E.'s would be very time consuming and would not add greatly to the significance of the results.

between the amounts of movement under these two conditions. We therefore conclude that when nervous tension is increased as by whispering prior to internal speaking, tongue movements occur more readily than they do when this tension is lacking.

TABLE IV. *A*

COMPARISON OF THE AMOUNT OF MOVEMENT DURING INTERNAL SPEECH FOLLOWING OVERT SPEECH, WITH THE AMOUNT DURING INTERNAL REPETITIONS OF THE SAME TASKS WITHOUT PRECEDING OVERT SPEECH

The comparison is made in terms of the percentage of cases in the first category which equaled or exceeded and the percentage which fell below the average of all cases in the second category. This same method of comparison is followed in Tables IV. *B* to VIII. inclusive.

Upper Curve				Lower Curve			
Above Average		Below Average		Above Average		Below Average	
No.	%	No.	%	No.	%	No.	%
25	83.4	5	16.6	23	88.4	3	11.6

TABLE IV. *B*

COMPARISON OF THE AMOUNT OF MOVEMENT WITHOUT INCREASED TENSION WITH THE AMOUNT OF MOVEMENT WITH INCREASED TENSION

Upper Curve				Lower Curve			
Above Average		Below Average		Above Average		Below Average	
No.	%	No.	%	No.	%	No.	%
6	12.0	44	88.0	2	4.0	48	96.0

The group V. situations, parts '1' and '3' (accompanied by distractions) compared with parts '2' and '4' (not accompanied by distractions) respectively, show the amount of the movement occurring during situations of severe distraction. The averages of about five hundred ratios are:

Averages for part 1..... 2.1 and 2.0
Averages for part 2..... 1.3 and 1.3
Difference..... .8 and .7

Table V. *A* shows that 92.5 per cent. and 90 per cent. of the ratios of part '1' are above the averages of part '2.'

Averages for part 3 ..... 1.9 and 1.8
Averages for part 4 ..... 1.5 and 1.3
Difference ..... .4 and .5

TABLE V. *A*

COMPARISON OF THE AMOUNT OF MOVEMENT WITH DISTRACTION WITH THE AMOUNT OF MOVEMENT WITHOUT DISTRACTION. (FOR THE TASK 'EXPERIMENTAL PSYCHOLOGY.')

Upper Curve				Lower Curve			
Above Average		Below Average		Above Average		Below Average	
No.	%	No.	%	No.	%	No.	%
37	92.5	3	7.5	36	90.0	10	10.0

TABLE V. *B*

COMPARISON OF THE AMOUNT OF MOVEMENT WITH DISTRACTION WITH THE AMOUNT OF MOVEMENT WITHOUT DISTRACTION. (FOR THE TASK, RECITING THE MULTIPLICATION TABLES.)

Upper Curve				Lower Curve			
Above Average		Below Average		Above Average		Below Average	
No.	%	No.	%	No.	%	No.	%
31	77.5	9	22.5	30	77.0	9	23.0

Table V. *B* discloses the fact that 77.5 per cent. and 77 per cent. of the ratios of part '3' are above the averages of part '4.' Both of the comparisons made here indicate that there is a significant difference between the conditions compared, particularly in the case of the first comparison. The conclusion follows that when laboring under a severe distraction, the subjects tend to greater movement of the tongue than when they are under lesser strain.

Another distraction was filing on a piece of brass while the subject performed a problem in mental multiplication (group V., part '6'). Mental multiplication was also carried out in the absence of any distraction (group I., parts '3' and '4'). The averages of approximately eight hundred ratios are:

Averages with distraction ..... 1.6 and 1.5
Averages without distraction ..... 1.2 and 1.2
Difference ..... .4 and .3

TABLE VI

COMPARISON OF THE AMOUNT OF MOVEMENT DURING MENTAL MULTIPLICATION WITH DISTRACTION WITH THE AMOUNT OF MOVEMENT DURING MENTAL MULTIPLICATION WITHOUT DISTRACTION

Upper Curve				Lower Curve			
Above Average		Below Average		Above Average		Below Average	
No.	%	No.	%	No.	%	No.	%
27	67.6	13	32.4	25	62.6	15	37.5

From Table VI. we find that 67.6 per cent. and 62.6 per cent. of the ratios with distraction are above the average of the ratios without distraction. This is probably somewhat significant, but the percentages are not high enough to warrant our drawing any definite conclusion as to the facilitating influence of the filing distraction on tongue movements.

To ascertain the effect of being pressed for speed in performing a given task, the ratios of the movements recorded while reciting the alphabet several times, as rapidly as possible with constant urge to increase the speed with each repetition (group V., part 5), were compared with the average ratios of movements during 'normal' internal speech such as group I. situations.

TABLE VII

COMPARISON OF THE AMOUNT OF MOVEMENT DURING EXTREMELY RAPID INTERNAL SPEECH WITH THE AMOUNT OF MOVEMENT DURING NORMAL INTERNAL SPEECH

Upper Curve				Lower Curve			
Above Average		Below Average		Above Average		Below Average	
No.	%	No.	%	No.	%	No.	%
26	65.0	14	35.0	22	55.0	18	45.0

Averages with speed	2.2 and 2.0
"Normal" averages	1.2 and 1.2
Difference	1.0 and .8

The absolute differences between these ratios are large, but when we look at Table VII. and find that only 65 per cent

and 55 per cent. of the ratios of movements while speeding are above the averages of the ratios of movements during 'normal' internal speech, we must conclude that these differences are not very significant. A glance at Table XIII. (Appendix) column '5,' will explain why the absolute difference is so great and yet not significant. Two subjects, 'B' and 'F,' practically whispered when reciting the alphabet under these conditions, and the rest of the subjects behaved very much as they did during 'normal' internal speech.

Having noted thus far that the increase in the amount of tongue movements is primarily determined by conditions of nervous tension, rather than the content of internal speech, it becomes of interest to ascertain how the amount of movement during the 'normal' internal speech (situations of group I.) compares with the amount of movement during blank situations.

TABLE VIII

COMPARISON OF THE AMOUNT OF MOVEMENT DURING NORMAL INTERNAL SPEECH WITH THE AMOUNT OF MOVEMENT DURING THE BLANK SITUATIONS

Upper Curve						Lower Curve					
Above Ave.		Equal		Below Ave.		Above Ave.		Equal		Below Ave.	
No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
23	45.2	14	27.4	14	27.4	16	32.0	21	42.0	13	26.0

Averages during 'normal' ..... 1.2 and 1.2
Averages during 'blank' ..... 1.1 and 1.1
Difference ..... .1 and .1

The absolute difference is slight, and a glance at Table VIII., showing 42.2 per cent. and 32 per cent. of the 'normal' ratios above and 27.4 per cent. and 42 per cent. equal to, and 27.4 per cent. and 26 per cent. below the averages of the 'blank' ratios, confirms the conclusion that the difference between the two groups of ratios is not very significant. Therefore tongue movements during internal speech without any increased tension are not significantly greater than those during 'blank' situations.

An attempt was made to get at the effect on tongue move-

ments of performing a task involving difficult verbal formulation. For this purpose average ratios of movement during writing from dictation and during the reading of a very difficult Polish paragraph were compared with the average ratios of 'normal' movements. The first group (writing) is exactly equal to the normal group. The average ratio of the second group (reading) is actually .05 less than the average ratio of the normal. This is surprising. However, the Polish proved to be so very difficult that most subjects did not even attempt pronunciation, but merely slurred it over. In this case it is to be expected that the Polish material would effect results more similar to the blank situations than to any other. In view of these facts, we must conclude that increasing the difficulty of the material of inner speech alone, does not increase the amount of tongue movements. It appears to have no effect.

To determine the effect of emotional disturbance upon parts of the tongue movements two methods were used: (1) during some experiment the subject was pinched, slapped in the face, hair pulled, or struck on the back, and (2) at the end of every sitting the subject was questioned as to his mood during the day. The first method showed no visible change either in the quality or quantity of movement, nor was any definite change noticeable when the subject himself reported emotional disturbance, except in two cases where there seemed to be slightly more tremor than ordinary.

In view of all of the above facts, it seems clear that the amount of tongue movement is not dependent upon the content of internal speech, nor even upon the presence or absence of internal speech, but upon other bodily activity which either precedes or accompanies internal speech.

Discussion

The results of this study have shown: (1) that there is no qualitative correspondence between overt speech and implicit speech during the verbal formulation of the same ideas, and further that there is no constant similarity between implicit movements made during this same verbal formulation; there-

fore the identity of the tongue movements and the elements of internal speech is precluded, (2) that a shift to other language mechanisms probably does not take place, and (3) that there is no evidence for quantitative variations in implicit speech corresponding to changes in internal speech.

Listing the various tasks in the experiment according to their rank in the accompanying amount of tongue movement we have:

1. Reciting silently the alphabet as rapidly as possible. Average 2.10.

2. Thinking 'experimental psychology' while singing 'ah' and tapping the table. Average 2.05.

3. Reciting silently the multiplication tables of 7's and 9's, while singing 'ah' and tapping with the fingers. Average 1.85.

4. Thinking following whispering. Average 1.75.

5. Mental multiplication during filing distraction. Average 1.55.

6. Reciting silently the multiplication tables of 7's and 9's after doing so under distraction. Average 1.40.

7. Thinking 'experimental psychology' after doing so under distraction. Average 1.30.

8.5. Writing a passage from dictation. Average 1.20.

8.5. Thinking without any distraction and following no distraction. Average 1.20.

10. Reading silently a Polish passage alternating with whispering of it. Average 1.15.

11. Blank. Average 1.10.

The total range of movement as shown in this table is from 1.10 to 2.10. The indication is that the amount of movement is proportional to the distraction, the amount of other bodily activity involved, and the amount of tension as shown in trying to think at high speed and the like. But to specify definitely which of the experimental tasks involve the greater and the lesser amounts of nervous tension is practically impossible in the present state of our knowledge. However it is possible to classify these ten tasks into two general groups: (1) the tasks involving the more difficult

problems, such as thinking under severe distraction, and having to intersperse each repetition of a thought word with the same word whispered, and (2) the situations involving just thinking without any simultaneous external or bodily activity. It is interesting to note that all situations falling into the first class, occupy the first five ranks in order of movement; all situations falling in the second class occupy the last five ranks; and also that the two situations in class two which most resemble those of class one, occupy ranks '6' and '7.'

The suggestion from this work is, then, that the overt movements of the tongue which are observed during internal speech ought to be interpreted as the mere chance result of neural irradiation, rather than as specific elements in the thought processes. In this respect they resemble drumming with the fingers, facial contortions, tics, and the like, which appear in individuals during emotional stress. This supports the position taken by Lashley ('24) in so far as increased irradiation goes with increased activity of the tongue, but does not support his statement that specific implicit movements are more likely to occur at such times.

To determine whether the tendency to such random motor activity varies consistently from one individual to another, the other activities of the subjects under tension were compared with the amount of tongue movement. Data on facial expression for six of the subjects were available from the experiments of Landis ('24), who measured the amount of muscular movement shown during emotional situations. The correlation obtained between the average amount of movement of the facial musculature in emotion, and the average amount of tongue movement in this experiment is $R = 0.37$. This would indicate that there is some positive relationship, but it is not high enough to be diagnostic. Furthermore it is based on only six cases. It is interesting to note, however, that the person ranking first (out of twenty-five subjects) in facial movements, also ranked first (out of eighteen subjects) in tongue movements.

On the whole, the results of these experiments throw no

light upon the actual mechanism of internal speech. There is no indication of a correspondence between movements of the tongue and verbal formulations in thought. This leaves only the hypothesis that the activities are intra-neural, and do not necessarily involve complete motor expression at each stage of the process.

Conclusions

1. Movements of the tongue are not universal in internal speech or verbal thought.

2. When they do occur, they correspond to movements in overt speech of the same words only in 4.4 per cent. of the cases of this experiment.

3. Repetition of the same verbal thought is accompanied by repetition of similar tongue movements in only 10 per cent. of the cases where movement occurs.

4. When tongue movements do not occur, it is not likely that other overt language mechanisms are functioning in place of the tongue.

5. Movements of the tongue occur more readily after the tongue has been activated in overt speech.

6. They occur more readily under distraction.

7. They occur more readily when working for speed.

8. Emotional disturbance has only a very slight effect on the quantitative increase of tongue movements, within the limits of this experiment.

9. Tongue movements are not so much dependent upon the content of internal speech as they are upon conditions of nervous irradiation and muscular tonus.

Bibliography

Adrian, E. D. The Conception of Nervous and Mental Energy (I). *Brit. J. Psychol.*, 1923, *14*, 121–125.

Bain, A. The Senses and the Intellect. New York, 1879.

Ballet. Le Langage Interieur et les Diverses Formes de l'Aphasie. Paris, 1886.

Bartlett, F. C., and Smith, E. M. Is Thinking Merely the Action of Language Mechanisms (I). *Brit. J. Psychol.*, 1920, *11*, 54–62.

Clark, R. S. An Experimental Study of Silent Thinking. *Arch. of Psychol.*, 1922, No. 48, IV., 96.

Courten, H. C. Involuntary Movements of the Tongue. *Yale Psychol. Studies*, 1902, *10*, 93–96.

CURTIS, H. S. Automatic Movements of the Larynx. *Amer. J. Psychol.*, 1899, *11*, 237–240.

EGGER. La Parole Interieure. 1881.

HANSEN AND LEHMAN. Über unwillkürliches Flüstern. Wundt's Phil. Stud., 1895, 11.

HEAD, H. The Conception of Nervous and Mental Energy (II). *Brit. J. Psychol.*, 1923, *14*, 126–147.

LANDIS, C. Studies of Emotional Reaction (II). General Behavior and Facial Expression. *J. Comp. Psychol.*, 1924, *4*, 447–501.

LASHLEY, K. S. The Neural Mechanism of Voluntary Movement. *Psychol. Bul.*, 1924, *21*, 94.

LEMAITRE, A. Observations sur le Langage Interieur des Enfants. *Arch. de Psychol.*, 1905, *4*, 1–43.

MYERS, C. S. The Conception of Nervous and Mental Energy (III). *Brit. J. Psychol.*, 1923, *14*, 148–152.

OTIS, A. S. Do We Think in Words? *Psychol. Rev.*, 1920, *27*, 399–419.

PAULHAN. Le Langage Interieure et la Pensée. *Revue Philosophique*, 1886, *21*.

PEAR, T. H. Is Thinking Merely the Action of Language Mechanisms (III). *Brit. J. Psychol.*, 1920, *11*, 71–80.

PIÉRON, H. Le Cerveau et la Pensée. Paris, 1923.

PINTNER, R. Inner Speech During Silent Reading. *Psychol. Rev.*, 1913, *20*, 129–153.

REED, H. B. The Existence and Function of Inner Speech in Thought Processes. *J. Exper. Psychol.*, 1916, *1*, 365–392.

RIBOT, T. Les Mouvements et leur Importance Psychologique. *Revue Philos.*, 1879, *8*.

RIBOT, T. Le Probleme de la Pensée sans Images et sans Mots. *Revue Philos.*, 1913, *76*, 50–68.

ROBINSON, A. Is Thinking Merely the Action of Language Mechanisms (IV). *Brit. J. Psychol.*, 1920, *11*, 81–86.

STRICKER. Studien über die Sprachvorstellungen. Wien., 1880.

THOMSON, G. H. Is Thinking Merely the Action of Language Mechanisms (II). *Brit. J. Psychol.*, 1920, *11*, 63–70.

WATSON, J. B. Behavior, An Introduction to Comparative Psychology. New York, 1914.

WATSON, J. B. Psychology from the Standpoint of the Behaviorist. Philadelphia, 1919.

WATSON, J. B. Is Thinking Merely the Action of Language Mechanisms (V). *Brit. J. Psychol.*, 1920, *11*, 86–104.

WYCZOIKOWSKI, A. Theoretical and Experimental Studies in the Mechanism of Speech. *Psychol. Rev.*, 1913, *20*, 448–458.

APPENDIX

Ratios of Amount of Movement by Subjects and Situations

TABLE IX

THE AVERAGE AMOUNT OF MOVEMENT FOR EACH SUBJECT ON EACH TASK IN THE GROUP I SITUATIONS

Upper Curve

Subjects	1[1]	2	3	4	5	6	7	Blank	Ave.
A	1.9	1.8	1.2	1.8	2.3	—	1.9	—	1.8
D	1.3	1.0	1.0	1.2	—	—	1.3	—	1.2
H	1.0	1.1	1.1	1.2	1.2	1.0	1.0	—	1.1
L	1.2	1.3	1.3	1.8	1.2	1.3	—	—	1.3
M	1.2	1.0	1.1	1.2	1.0	1.0	1.1	1.0	1.1
O	1.1	1.2	1.1	1.1	1.1	1.1	1.1	—	1.1
P	1.2	1.1	1.4	1.1	1.1	1.2	—	1.2	1.2
Q	1.0	1.1	1.0	1.0	1.0	1.0	—	1.1	1.0
Ave.	1.2	1.2	1.1	1.3	1.3	1.1	1.3	1.1	1.2[2]

Lower Curve

Subjects	1	2	3	4	5	6	7	Blank	Ave.
A	1.4	—	1.1	1.6	2.1	1.6	1.6	—	1.6
D	1.1	1.1	1.0	1.2	—	—	1.2	—	1.1
H	1.0	1.1	1.1	1.1	1.0	1.0	1.0	—	1.0
L	1.1	1.3	1.3	1.7	1.5	1.5	—	—	1.4
M	1.1	1.0	1.1	1.2	1.0	1.0	1.1	1.0	1.1
O	1.1	1.1	1.1	1.1	1.3	1.1	1.1	—	1.1
P	1.3	1.1	1.4	1.1	1.1	1.1	—	—	1.2
Q	1.0	1.1	1.0	1.0	1.0	1.0	—	1.1	1.0
Ave.	1.1	1.1	1.1	1.3	1.3	1.2	1.2	1.1	1.2

[1] The numerals at the head of each column refer to the number of the tasks as numbered on page 11.

[2] Blanks not included in the average. This is true of Tables IX.–XIII. inclusive.

TABLE X

THE AVERAGE AMOUNT OF MOVEMENT FOR EACH SUBJECT ON EACH TASK IN THE GROUP II SITUATIONS

Upper Curve

Subjects	1[1]	2	3	Blank	5	6	7	8	9	Ave.
D.....	1.4	2.1	1.5	1.2	1.3	2.7	3.3	1.2	—	1.8
H.....	1.1	1.5	1.3	1.1	2.2	1.8	2.5	—	—	1.7
K.....	1.8	2.2	1.1	1.2	—	—	—	—	—	1.7
L.....	1.4	—	1.1	1.1	1.2	1.2	1.3	—	—	1.2
M.....	1.1	1.2	1.1	1.1	1.2	1.1	—	—	—	1.1
B.....	—	—	—	—	—	—	—	4.5	—	4.5
G.....	1.1	—	—	—	—	—	—	1.2	—	1.2
N.....	1.9	—	—	—	—	—	—	1.7	—	1.8
C.....	—	—	—	—	—	—	—	—	1.3	1.3
F.....	—	—	—	—	—	—	—	—	2.9	2.9
R.....	—	—	—	—	—	—	—	—	1.3	1.3
Ave.....	1.4	1.7	1.2	1.1	1.5	1.7	2.3	2.2	1.8	1.8[2]

Lower Curve

Subjects	1	2	3	Blank	5	6	7	8	9	Ave.
D.....	1.6	1.4	1.3	1.1	1.3	3.4	2.0	—	—	1.8
H.....	1.1	1.2	—	1.1	2.0	2.9	1.8	—	—	1.8
K.....	1.7	2.1	1.1	1.1	—	—	—	—	—	1.6
L.....	1.4	—	1.1	1.1	1.2	1.2	1.4	—	—	1.3
M.....	1.1	1.1	1.1	1.1	1.2	1.2	—	—	—	1.1
B.....	—	—	—	—	—	—	—	3.2	—	3.2
G.....	—	—	—	—	—	—	—	—	—	—
N.....	1.3	—	—	—	—	—	—	1.5	—	1.4
C.....	—	—	—	—	—	—	—	—	1.4	1.4
F.....	—	—	—	—	—	—	—	—	1.9	1.9
R.....	—	—	—	—	—	—	—	—	1.8	1.8
Ave.....	1.4	1.4	1.2	1.1	1.4	2.2	1.7	2.4	1.7	1.7

[1] The numerals at the head of each column refer to the number of the tasks as numbered on page 12.

[2] No. 3 not included in this average.

TABLE XI

THE AVERAGE AMOUNT OF MOVEMENT FOR EACH SUBJECT ON THE TASKS OF GROUP III

Subjects	Upper Curve		Lower Curve	
	Internal Speech	Blank	Internal Speech	Blank
D.....	1.5	3.0[1]	1.2	2.7[1]
G.....	1.1	1.1	1.1	1.0
H.....	1.2	—	1.3	—
M.....	1.0	1.0	1.0	1.0
Q.....	1.0	1.1	1.1	1.1
Ave.....	1.2	1.1	1.1	1.0

[1] Thrown out because the subject did not follow the instructions of the experimenter.

TABLE XII

THE AVERAGE AMOUNT OF MOVEMENT FOR EACH SUBJECT AT EACH SITTING ON THE VARIOUS TASKS OF GROUP IV

Subjects	Upper Curve				Lower Curve			
	Silent Reading	Writing Dict.	Blank	Aver.	Silent Reading	Writing Dict.	Blank	Aver.
C.....	1.1	1.1	1.1	1.1	1.1	1.0	1.1	1.0
C.....	—	1.5	—	1.5	—	1.4	—	1.4
D.....	—	1.1	1.0	1.1	—	1.1	1.0	1.1
D.....	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
E.....	—	1.4	—	1.4	—	1.2	—	1.2
F.....	—	1.3	—	1.3	—	1.1	—	1.1
G.....	—	1.1	—	1.1	—	1.1	—	1.1
G.....	1.0	1.1	—	1.1	1.0	1.0	—	1.0
H.....	—	1.9	—	1.9	—	1.0	—	1.0
J.....	1.3	1.2	—	1.3	1.2	1.3	—	1.3
J.....	1.2	1.2	1.1	1.2	1.1	1.1	1.1	1.1
O.....	—	1.3	—	1.3	—	1.2	—	1.2
R.....	—	1.8	1.4	1.8	—	1.4	1.4	1.4
Ave..	1.1	1.3	1.1	1.3	1.1	1.1	1.1	1.1

TABLE XIII

THE AVERAGE AMOUNT OF MOVEMENT FOR EACH SUBJECT ON EACH TASK IN THE GROUP V SITUATIONS

Upper Curve

Subjects	1[3]	2	3	4	5	6	Blank	Ave.
A[1]	3.4	1.8	2.5	2.0	1.7	1.8	1.5	2.2
B[1]	1.7	1.5	1.5	1.9	4.3	1.9	1.4	2.1
C[1]	2.4	1.3	2.1	1.3	1.2	1.1	1.3	1.6
D[1]	1.7	1.1	1.7	1.1	1.1	1.3	1.1	1.3
E[2]	2.1	1.4	2.1	1.7	1.7	1.1	1.0	1.7
F[2]	2.4	1.3	2.2	1.5	6.9	4.0	1.4	3.1
G[2]	1.6	1.1	1.6	1.3	1.1	1.3	1.1	1.3
H[1]	1.5	1.1	1.4	1.1	1.2	1.2	1.1	1.2
I[2]	2.0	1.4	2.2	1.6	1.2	1.3	1.0	1.6
J[2]	2.5	1.3	1.9	1.2	1.3	1.3	1.1	1.6
Ave.	2.1	1.3	1.9	1.5	2.2	1.6	1.2	1.77

Lower Curve

Subjects	1	2	3	4	5	6	Blank	Ave.
A	2.6	1.5	2.0	1.5	1.4	1.7	1.3	1.8
B	1.6	1.5	1.5	1.6	3.4	1.5	1.2	1.9
C	2.6	1.3	1.7	1.1	1.2	1.1	1.3	1.5
D	2.2	1.2	2.0	1.2	1.2	1.3	1.1	1.5
E	1.9	1.3	1.6	1.4	1.4	1.1	1.0	1.5
F	2.2	1.3	1.9	1.2	7.1	3.2	1.4	2.8
G	1.4	1.1	1.6	1.3	1.0	1.2	1.1	1.3
H	1.3	1.1	1.2	1.1	1.1	1.1	1.1	1.1
I	1.7	1.4	2.0	1.6	1.1	1.2	1.0	1.5
J	2.4	1.4	2.1	1.2	1.4	1.4	1.2	1.6
Ave.	2.0	1.3	1.8	1.3	2.0	1.5	1.2	1.65

[1] Averages of performances on six different days.

[2] Averages of performances on two different days.

[3] The numerals at the head of each column refer to the number of the tasks as numbered on page 12.

FATIGUE TESTS AND INCENTIVES[1]

BY HELEN FRANCIS WHITING AND HORACE BIDWELL ENGLISH

Psychology Laboratory, Wellesley College

The practical need for simple and accurate tests of fatigue is one that the psychologist may not ignore. This need, accordingly, constitutes the starting point of these investigations. Although there is considerable agreement that fatigue may not be defined wholly in terms of lessened efficiency for mental or physical work, it is this lessened efficiency which gives to the problem its practical interest and all fatigue tests have proceeded by measuring this work-decrement. A preliminary analysis indicates that this work-decrement may show itself in two ways: loss of speed and loss of accuracy. Cutting somewhat across these two is the relation of fatigue to difficulty. It is possible that while fatigue may not affect work of little difficulty at all, it may greatly influence more difficult tasks; that is it may raise the work threshold for either accuracy or speed.

Corresponding to this analysis, six simple tests were devised for Part I., three of a physical nature, and three of a mental nature,—each one to test the effect of fatigue upon only one of the three elements: accuracy, difficulty, speed. The procedure was as follows: the subjects, sixteen undergraduates of Wellesley College, reported for testing at or before eight-thirty in the morning. They then followed their normal routine as college students, reporting for another test at four-thirty in the afternoon. Each student was tested on

[1] The experiments herein described were carried out by H. F. Whiting; those of Part I. under the general direction of L. R. Geissler, those of Part II. under the general direction of H. B. English. They formed the basis of a dissertation offered to the Faculty of Wellesley College by Miss Whiting (now Mrs. R. G. Goesle) in partial fulfilment of the requirements for the Master of Arts degree. The dissertation was condensed for publication by Dr. English who is thus responsible for the present form, though Miss Whiting's treatment has been left unaltered in many places. Since this paper was prepared for publication, B. Muscio has published two papers on fatigue which take a similar position. (See *Brit. J. Psychol.*, 13, 31–46 and 150–162.)

from three to five days. Every precaution was taken to insure practical equivalence of difficulty in the tests, where they were not in fact identical, night and morning. Opportunity is thus afforded to compare the loss of efficiency caused by the normal fatigue of the day, since the subjects were virtually not fatigued at all when they first reported, and had done the major part of their normal day's work when they came for the second testing. Subjective estimations of fatigue conditions were recorded at each sitting. The experiments of Part I. were carried out in the academic year 1918–1919.

In Test 1 an attempt is made to test the effect of fatigue upon the accuracy of physical work by considering its effect upon movement of the arm. The blindfolded subject, drew her pencil along a straight edge for a standard length as determined by a stop, and immediately reproduced this length with the stop removed. A simple device enabled the experimenter to record graphically the exact length of line reproduced. The three standard lengths, which were never seen by the subject, were, in the order of presentation, 7.4 cm., 10 cm., and 12.6 cm. The failure of the capacity here measured, *viz.*, to make a simple movement accurately, is responsible for many accidents. It has been unhesitatingly ascribed to fatigue in many studies of industrial accidents and their relation to fatigue.

Test 2 was devised to measure the effect of fatigue upon mental accuracy. The subject was required to divide four lines, measuring 8.3, 15.3, 14.2 and 9.0 cm., into halves and thirds by simple occular estimation. The lines were shown singly through a narrow window. It is not so obvious that this test is directly related to practical conditions. To some extent the problem is bound up with the stale old question of general *vs.* special abilities. Yet not only is accurate spatial judgment required in many practical situations, but its failure may again, as in the case of machine tenders, be at least a contributory cause of accident.

Test 3 was designed to indicate the effect of fatigue upon speed of mental work. For this the familiar addition tests

seemed most available. The particular form used was as follows: A column of 3's and 4's in chance order was placed before the subject. These were added cumulatively, one at a time, to a number varying with each test which was given as a starting point. This was continued for thirty minutes. The instructions stressed speed. Two such series were given in each experimental period. It is generally held that continuous addition, with instruction for speed, tests some such general function—one ought to say faculty, even though the term is in bad odor—as concentration, upon which speed of mental operations in general is contingent. If this is true, the practical value of the test is apparent.

Test 4 was included as a test of the effect of fatigue upon the difficulty of physical work. The subject was required to hold a metallic pencil with unsupported arm for thirty seconds in each of three consecutive holes of the Whipple Steadiness Tester. The holes were 11, 10, and 9/64 of an inch in diameter. Contacts were electromagnetically recorded on the kymograph. Each hand was tested for the three holes in succession, the right hand first. The value and even necessity of steadiness in many mechanical pursuits is apparent. And steadiness of hand has long been considered an especially good diagnostic test of neuro-muscular condition.[1]

The fifth test was planned to disclose the effect of fatigue on the difficulty of mental work. Ten multiplication problems of increasing difficulty were successively presented through a small window in a screen and the subject was required to solve them without the aid of pencil and paper. No two problems were the same, but every effort was made to keep corresponding problems of equivalent difficulty. It was at first planned to allow unlimited time as the test was only for difficulty of mental work and not for speed, but one of the subjects spent so much time that it was found necessary to limit the time to fifteen minutes. All the other subjects were able in most cases to finish within this limit.

[1] To one of the writers, however, it seems that the test frequently tests nothing but voluntary attention. The holes are large enough so that a normal and steady individual can avoid contacts as long as he 'can keep his mind on it.' The simplicity of the situation however, makes this by no means easy.

In the sixth test, the subject was required to tap as fast as possible upon a metal plate with a metal stylus. Contacts were electromagnetically recorded upon the kymograph. The two periods of thirty seconds each for each hand were separated by the period given to the other hand. The test was obviously intended to measure the speed of physical work. This test, alone of the six, itself excites a marked feeling of neuro-muscular fatigue.

The results with test 1 show little if any significant difference between morning and afternoon. We may compare the "reproduced line" of morning and afternoon with the standard and also the two reproduced lines with each other. In 50 out of 90 cases the reproduced line was shorter in the afternoon than the corresponding line in the morning. Treating the results from a slightly different standpoint, the reproduced line was longer than the standard in 74 morning cases and in 61 afternoon cases, was shorter than the standard in 84 morning cases and in 96 afternoon cases. There is thus a fairly well marked tendency to overestimation of the movement in the afternoon. On the other hand, the average error and the mean variation are slightly less in the afternoon. (9.6 mm. and 6.1 mm. morning and 9.3 mm. and 5.0 mm. in the afternoon; obviously hardly a significant difference.) In 40 cases the error was greater in the afternoon than in the morning of the same day, in 46 cases it was less. Of the 16 subjects, 8 showed a tendency to average slightly better in the afternoon, 7 in the morning, one did equally well. Six were more variable in the afternoon, 10 were less.[1]

Very similar results were found in test 2. In 49 cases the afternoon performance was worse, in 44 it was equal, in 47 it was better. Accuracy of spatial estimation is unaffected by the normal daily fatigue of these subjects.

In test 3 (addition) it will be remembered that the emphasis was upon speed. As a matter of fact, there was very little inaccuracy and almost never was the final sum more than one digit wrong. In 31 out of 108 cases, the accuracy was less in

[1] The complete data fill many pages; they may be consulted in the Wellesley College Library and in the Wellesley Psychological Laboratory.

the afternoon, in 60 it was equal, and in 17 cases it was greater. Where the subjects reported themselves as feeling more fatigued in the afternoon, the results are somewhat more favorable to that period; in 38 cases they do equally well, and the number of cases in which they do better or worse in the afternoon is equal at 16. Thus as regards accuracy of addition, there is a clear cut negative correlation between subjective estimation of condition and objective results.

Disregarding accuracy, we find that in 66 cases the speed is greater in the afternoon, in 12 it is equal, in 30 it is less. Where the subjects report themselves as more fatigued in the afternoon, the proportions are almost exactly the same, the number of cases being 40, 9, and 21 respectively. The average number of items added per trial was 19.5 in the morning and 21.0 in the afternoon. The m.v. was 5.0 in both cases. Clearly then unless one is required to add the column with infallible accuracy in one trial, one had better make the attempt after a day's work. Even if scrupulous accuracy is needed, the most efficient method would seem to be to add late in the day but to check the results.

Steadiness tells the same story. In 57 cases more contacts were made in the afternoon, in 26 cases an equal number, in 85 cases a less number. The situation is not altered by considering right and left hands separately. The variability of the performances is such as to deprive the means of all significance; in one set of data (with Hole 5) the ratio of variability exceeds unity.

The particular purpose of this test however, was to measure the effect of fatigue upon difficulty. Hole 5 is a fairly large hole, Hole 7 is small. Thus one might actually do better on the comparatively easy task but fall down relatively either to the morning's performance with the same hole or to the afternoon's performance with the easier task. Neither of these results follows. The numbers of cases where the performance was worse, equal, or better in the afternoon are 18, 4, and 32 respectively, or about the same proportion as when all three holes are considered. A failure relative to the afternoon's performance with the easier task would be

evidenced by an increase in the ratio between Hole 5 and Hole 7 as compared with the morning's performance. Again no such difference was found. On the contrary, a small decrease was found in a slight majority of the cases. We may conclude that under our conditions, steadiness is increased in the afternoon, and that if anything, this gain in the afternoon is greater in the harder tasks.

Little more satisfaction is to be had with the test for mental difficulty—mental multiplication. In a majority of cases the subjects finished very nearly the same number of problems night and morning, thus showing that they were observing the instructions which stressed accuracy rather than speed. Unfortunately the accuracy was also virtually equal. Using an arbitrary method of combining the results for accuracy and speed,[1] the score was smaller in 24 cases in the afternoon and greater in 23. The average score was 81.3 in the morning, 80.3 in the afternoon, the mean variations are 22.7 and 23.8 points respectively.

There remains the tapping test for speed in physical work. The story is quite consistent with the foregoing and may be briefly told. There was a slight but insignificant increase in the average number of contacts made in the afternoon.

Is the initial assumption that the subjects are fatigued by the work of the day invalid? They do not think so. Their subjective estimates of physical and mental fitness sometimes show, to be sure, that they considered themselves in as good or in better condition in the afternoon, but this is not the usual case. The above figures are not altered substantially if we eliminate all cases where the subject did not feel fatigued at the second testing. Indeed in some cases, the performances were slightly better under these circumstances. The index of correlation between subjective estimation of condition and performance is negative but too low to be valid.

The unanimity with which the tests fail to disclose fatigue

[1] Suggested by Dr. Geissler; since the problems were of graded difficulty, higher credits were given for the solution of the more difficult. The number of cases is too small to permit of statistical determination of the best possible scoring formula but it is clear that the plan adopted secures substantial justice.

is suggestive almost equally with positive results. It is hard to believe that no fatigue was present. The normal routine of a college undergraduate is not intended to be excessively fatiguing, and judging by the academic results ordinarily attained, probably is not. But merely keeping awake, walking about, and even partly listening to lectures must induce some fatigue, and as most of these students were of high standing, more than this was true of them. It is not contended that the tests are extraordinarily delicate, yet the scoring is sufficiently fine to record even quite moderate amounts of work-decrement. Yet in spite of the fact that there must be some fatigue, we find no measurable work-decrement and in some cases we find work-increment. Our data are too few to enable us to say with certainty what the true explanation of this somewhat anomalous state of affairs is. But the very uniformity suggests the operation of some constant factor.

We believe that this factor is the 'test attitude.' The attempt to make a good record is certainly a very constant factor in these tests; the subjects exhibited great interest in their records as well as in the records of their fellows. They knew that the experiments dealt with fatigue, and there were frequent indications that knowing themselves to be fatigued, they threw themselves into the game with increased energy. Such spurts are familiar factors in all efficiency measurements; they can only be eliminated by prolonging the tests beyond the possibility of spurting.

This hypothesis, therefore, formed the basis of the tests in Part II. We must make the testing period so long that any tendency to overcome the effect of fatigue by spurting will be itself overcome. A further line of investigation grew out of the same hypothesis: if fatigue can be overcome (temporarily) by the "test incentive," may it not be overcome by other incentives? One may seldom be sure that one has ruled out an incentive, but it is often possible to regulate conditions by determining that certain incentives shall be operative.

The experiments of Part II. took place in the academic

year 1919–20. There was a complete change of subjects: two, Dr. Lucy Wilson and Dr. H. B. English, were instructors in psychology; two, Miss D. Collier and Miss R. Prouty were graduate students in psychology. For the two former, 22 sittings were held; for the two latter, 25. The subjects worked without intermission for 90 minutes on eight tests arranged in what may be called an ascending and descending order as follows:

1. Reproduction of lines.
2. Division of lines.
3. Steadiness—right hand, three holes; then left hand, three holes.
4. Addition. 5 minutes.
5. Cancellation. 5 minutes.
6. Multiplication. 5 minutes.
7. Memorizing nonsense syllables. 15 minutes.
8. Cancellation, second trial. 5 minutes.
9. Speed of movement test.
   Right hand 60 seconds.
   Left hand 60 seconds.
   Right hand 60 seconds.
   Left hand 60 seconds.
10. Cancellation, third trial. 5 minutes.
11. Memorizing nonsense syllables, second trial. 15 minutes.
12. Multiplication, second trial. 5 minutes.
13. Cancellation, fourth trial. 5 minutes.
14. Addition, second trial. 5 minutes.
15. Steadiness test. As above.
16. Division of lines. As above.
17. Reproduction of lines. As above.

These tests may be described as follows. The first three are identical with three tests in Part I. In the addition test, the subject was given a starting number to which he added the digits cumulatively up to and including 9 and then repeated still cumulatively. Thus the series of numbers starting with 18 would be 19 21 24 28 33 39 46 54 63 64 66 69 etc. These sums were written by the subject with closed eyes on a large sheet of paper.

Cancellation is a new test which combines the tasks of the familiar cancellation test and the task of alphabet sorting.[1] Letters and numbers were typed on sheets in pied order but in lines. The task was to cancel first a letter, then a number, then the next letter in alphabetic order, then the next number and so on. If the subject finished two sheets, he was required to change the order and cancel first a number, then a letter.

In multiplication the multiplicands were presented to the subject on a small card. These were to be multiplied in order by each of four one place multipliers which were announced to the subject at the beginning of the test. The product was to be written down as a whole, not first the units place, then the tens place.

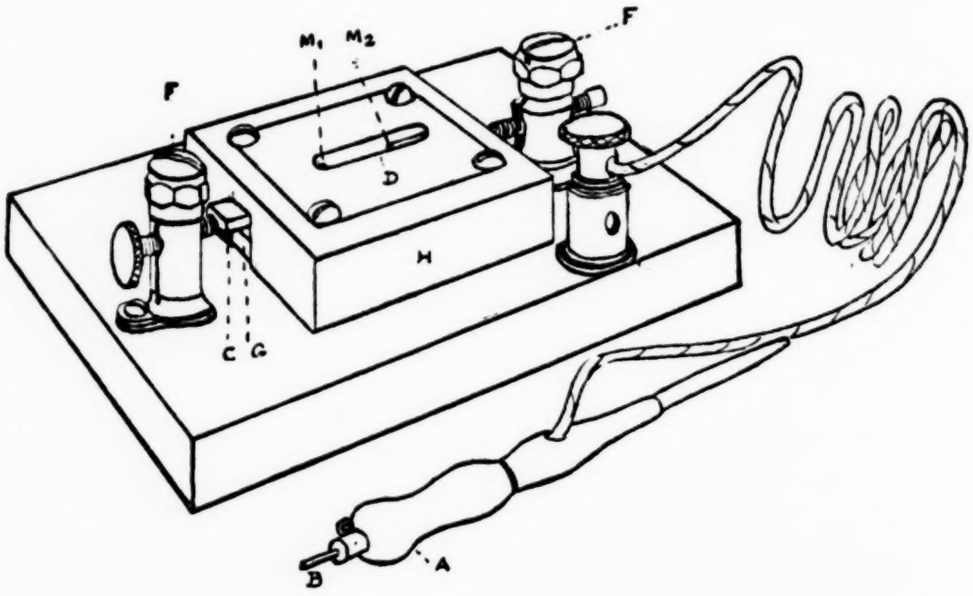
In learning nonsense syllables, the method of complete memorizing was used. Twelve syllables were used to a series; after the final and triumphant repetition of a series a new one was inexorably begun with the passage of one beat of the metronome. The rate of repetition was one syllable to each second. The usual rules for this method were faithfully followed, and nonsense syllables were constructed according to the rules formulated by Prof. E. A. McC. Gamble.[2]

The speed of movement test was devised by Dr. English as an improvement on tapping. Two criticisms are made of tapping: it is difficult to insure accurate registration, especially in view of the great variations in the method of tapping which the test inevitably permits; and secondly, these variations in method are so great as to make it impossible to ensure that you are really dealing with a single function. As subjects become tired, they bring new muscles into play. Both of these faults are obviated in the instrument used in this test. (See Fig. 1.) The sliding metal bar, *C*, makes and breaks an electric current on contact with the posts, *FF*. *C* slides easily over a smooth, oiled brass surface, *G*, passing back and forth through the ends of the block, *H*. It is moved by the

[1] The test was devised by Dr. English on the basis of his experience with alphabet sorting. *Cf.* 'The Intelligence of School Children,' *Psychol. Monog.*, 1917, 23, 266–331.

[2] *Cf. Psychol. Monog.*, 1916, 23, pp. 110–111 (Wellesley College Studies in Psychology, No. 2).

point of a stylus, *B*, which is drawn back and forth through a narrow slit in the brass plate, *D*, pressing alternately against the surfaces, M_1 and M_2 in *C*. The screws in *FF* are adjustable so that the exact length of the excursion of *B* may be regulated. The slit, *D*, is so narrow that only two types of movement are possible: either a straight forearm movement or a simple wrist movement. These two movements require a different position of the body and particularly of the arm to the instrument and are readily distinguished. They cannot be unconsciously substituted for each other. In both cases,



the muscles involved are relatively numerous so that nothing analogous to a nerve-muscle preparation is possible, but the muscles involved are constant. This device thus avoids the second objection to tapping. As for the first, accurate registration of the number of distinct movements made is insured by the break between M_1 and M_2. The stylus must traverse a distance of 1 cm. between each 'make' of the current and its break, one each side. This gives ample time for even the weak current necessary for most recording instruments to overcome the resistance at the contact surface. *C* is constantly in connection with one pole of the battery through its contact with *M*; either one or both posts, *FF* may be used. In this investigation, contacts were recorded from

only one side of the instrument and were recorded on the kymograph.[1] The bulge, *A*, in the stylus handle affords a firmer grip. This instrument was not devised until the experiment had been under way for some time; prior to that time, the subjects made a back and forth movement between the two sides of a Whipple Steadiness Lane (using the larger end). This fulfilled much the same function in fatiguing them but was not recorded or scored. The subjects were not aware of this, however, and made every effort to move as fast as possible.

A careful layout of the necessary material, and preparation on the part of both the experimenter and the subjects made it possible to run through the ninety minutes of these tests, passing from task to task, without intermission, with the loss of not more than ninety seconds. No opportunity was thus given the subject to recover from his exertion if he spurted in any tests without the effect showing in that or some other test. What are the objective results?

Fatigue, if it is to be measured in terms of work-decrement, by these tests, must show itself by a poorer performance in the second trial of the particular test compared with the first. In the reproduction of lines the error was greater in 86 cases, less in 77 cases, equal in 8. In amount of error, therefore, there seems scarcely any significant difference. In 102 cases, the reproduced line was shorter the second time than it was the first, while in 65 cases it was longer and in 4 equal. Here we seem to have a statistically significant difference, presumably of fatigue. On the other hand, where in Part I. the subjects showed a distinct tendency to draw the line shorter than the standard, they now show a distinct tendency to draw the line longer, though this tendency is less in the second trial.

In the division of lines test, we have no significant difference. In 99 cases there is improvement in the second trial, in 101 there is deterioration, in 56 there is neither. The results for single subjects are virtually the same. Neither in

[1] The Ewald chronoscope suggests itself as a convenient method of recording these movements in all cases where only totals are desired; it may also be used in connection with the graphic record.

range of error, mode, median, mean or mean variation are significant differences to be found between the first testing and the second.

With the steadiness tester, the subjects do better work the second time. It seemed to the subjects that this was at least partly due to the fact that just before the first test, the subject has been more or less active, at least to the extent of having walked some little way to reach the laboratory; before the second test there had intervened a period of 75 minutes of almost complete physical quiescence, broken only by the speed of movement test in the middle. One subject, however, showed a distinct deterioration in her second trial. This subject was extremely nervous throughout the tests, had much the largest number of contacts at all times, and distinctly tended to go to pieces under strain.

No new light comes with the results of the addition tests. There is a slight tendency to improvement the second time, probably statistically valid but so slight that it may be plausibly referred either to increasing adjustment to the test at its repetition, (which is the testimony of at least one subject) or to practice. The same is true of cancellation save that the same subject mentioned in connection with steadiness again shows deterioration where the others show slight improvement.[1] In multiplication there is a little improvement for all subjects. In 50 cases there is a gain on the second trial and in 41 there is a loss. This is paralleled by memorizing where the number of cases showing improvement is 48 compared with 43 showing a loss. (In both these cases, some uncertain results are omitted.)

Only in the speed of movement test do we get any marked deterioration. Here the results are clear cut. In only 6 cases was there any improvement in the second trial, while in 91 cases there was loss in speed. It should be noticed that the first and second testings were directly consecutive and that this test alone required rapid and severe physical exertion.

We had hoped to be able to measure the effect of various incentives in overcoming the work-decrement effect of fatigue.

[1] The first and fourth trials were compared in the cancellation test.

As our tests showed no such decrement, this is impossible. We may summarize briefly the effect of our incentives upon general test efficiency. For 15 sittings, the subjects had only such knowledge of their results as was inevitable from the nature of the tests. Beginning with the 16th they were told prior to beginning work what their previous results had been. In a third set of experiments, the subjects were divided into two competing pairs—W. with E. and P. with C.—and an appeal made to the spirit of rivalry, by informing each subject prior to beginning the test of his rival's results as compared with his own. All the subjects felt the genuineness of these incentives, especially of rivalry and seemed to themselves to be improving enormously in consequence. The results hardly bear this out. Work with knowledge certainly tends to reverse the direction of error in the lines tests but it does not influence its amount. And in some tests, there was a tendency to speed up at the expense of accuracy. There was a steady improvement in some tests, to be sure, but this improvement looks suspiciously like an extension into the second and third periods of the normal effect of practice, which had certainly not ceased in the sittings immediately prior. We are not convinced that the incentives were not operative; the subjects find it impossible to believe that the increased will and effort to win were entirely ineffective. But there is no objective proof for this belief; the efficiency curve is the normal practice curve.[1]

The reader may be assured that every effort has been made to find valid differences between the more and the less fatigued states of our subjects. No one likes to content himself with negative results. But mathematical treatment cannot make data yield significance they do not possess. In the case of our tests, whether we deal with individual measurements, with individual subjects, or with general averages, with net efficiency, with speed or with accuracy, or with variability, we are forced to conclude that there is virtually no real difference.

What are we to make of such results? The writers cannot

[1] Of course a plateau might have been imminent and been averted. The whole thing reverts to the status of speculation.

accept the view that the reason there were no fatigue 'effects' was because there was no fatigue. We do not believe that there is any plausible meaning that one can give to the term fatigue which would exclude the effect upon normal people of a day's work, however light. We should expect some fatigue, even if not a great amount. As for the experiments in Part II., the test was given late in the afternoon, and in the case of W. and E. usually followed the heaviest academic program of the week. While the experience of Part I. did not lead us to expect that this normal fatigue would by itself show in the test results, it was felt that it would add itself to the fatigue inherent in the tests and thus act something like the traditional straw that broke the camel's back. In point of fact, it seemed to the subjects that the tests were quite enough for that purpose. They were unanimously voted a ghastly combination. One subject complained of almost complete inability to do any further mental work after undergoing the torture and others spoke in like vein. It was not sheer boredom as might have been alleged had the subjects been condemned to ninety minutes work on one test, such as mental arithmetic, for instance. The tests were moderately interesting and the task was constantly shifting. If ever there was 'subjective' [1] fatigue, it was excited by this team of tests.

The next easiest solution is that the tests are bad tests for fatigue. Again we cannot accept this conclusion—at least without a struggle. The essential characteristic of a fatigue test is that it shall call upon one's mental or physical powers. It need not be, indeed for many reasons it would better not be, in itself fatiguing. Now every one of the tests here used definitely meets this criterion. If the attack is to be made on the tests themselves, it must center on their unreliability as means of measuring anything. Obviously the results would be vitiated if the mean variation were so great that any fatigue effects were liable to be swamped by 'accidental'

[1] No one with a behavioristic bent need be exercised over this 'archaic' term. For a long time psychologists of opposite camps will have to bear with each other in the matter of terminology. The writers would disavow any *exclusively subjective interpretation* of the term subjective. "Chacun a son gôut."

variations in the opposite direction. We need not examine this type of criticism at any length; all of the tests are familiar tests or adaptations of familiar tests, the value of which as such is reasonably established. In one test, to be sure, we reported that the ratio of variability was unusually high, but this was the result of lumping together the results from a small number of individuals differing greatly *inter se*. The individual variability is not too great in any single test to make it usable.

The lack of more extended investigation, as well as the negative nature of our results do not warrant us in putting forward a positive theory. Nevertheless, we propose to sketch an hypothesis, not entirely new, which is consistent with our facts. In order to state this hypothesis as briefly as possible, we shall omit all reference to the detailed facts from our own experiments and from those of others which suggest it. We do this the more willingly as we wish to be very clear that we do not claim even approximate factual verification for this hypothesis. It is intended as a challenge to further work, not as a summary of work done.

We must clearly distinguish between exhaustion and fatigue. The paradigm of exhaustion is the nerve-muscle preparation or the direct exhaustion of a muscle by electric stimulation. The work decrement of exhaustion follows the familiar 'Treppe' curve. Under non-experimental conditions this work-decrement may not be obvious, because of the operation of factors resulting in increased efficiency, but the total efficiency in such cases must be considered as the resultant of the two or more divergent factors—*i.e.*, partial exhaustion always results in a lessened capacity for work.

While exhaustion is thus primarily a physiological or specific and local phenomenon, we believe that fatigue is chiefly a 'subjective' phenomenon. Like exhaustion, it is a function of exertion. But unlike exhaustion, fatigue is not symptomatic of inability to do work. The relation of fatigue to work is more complex.

Considered as a subjective phenomenon, fatigue is seen to be a feeling of a complex sort closely akin to the appetites.

Its strong emotional coloring, the unanalyzable complex of visceral and organic sensations, the internal stimulus, and especially the direct and obvious relation to motor activity all suggest classification as an emotional appetite. As such, fatigue functions as a conscious (and negative) motive for action. Thus where exhaustion affects primarily the mechanism by which work is accomplished, fatigue tends to withdraw or reduce the motive power or drive. Or rather, the effect of fatigue is to raise the threshold at which a work motive may be effective, but does not lower the efficiency of the work, granted the adequacy of the positive motive.

This all-or-none effect of fatigue is concealed by two factors. We have in the first place the gradual lowering of efficiency due to increasing exhaustion. In the second place, while fatigue acts on the all-or-none principle at any given level, the fatigue itself may be of any degree from zero to a kind of torpor or stupor and thus be differentially 'effective' against the positive work motives.

In the economy of the organism, it is the function of fatigue to preserve the body from exhaustion of which, as we have said, it is a concomitant. It serves as a more direct and immediate warning of approaching exhaustion than work-decrement and in the absence of powerful motives for continued exertion acts as a direct inhibiter. We may hazard the guess that its physiological basis is some general toxemia (whether of general or local origin) affecting primarily the cerebrum.

Some of the puzzling results of these and other experiments obviously are easily explained under this hypothesis. The effects of physiological exhaustion as we have defined the term are too slight under the conditions of our experiments to be measurable. Moreover the test situation simply introduces a motive strong enough to counteract the effects of the fatigue motive, or rather too powerful to be affected by it at all. The two motives exist side by side but only one controls, the desire to do well on the tests. Hence we get not only no work-decrement correlative with the fatigue but no work-decrement at all.

Summary

1. Tests for speed, accuracy, and difficulty of both mental and physical work were given to 16 Wellesley College undergraduates in the morning before college work and in the afternoon following college work. No significant differences in efficiency were found.

2. A team of tests extending over 45 minutes of continuous work was followed without intermission by the same tests in reverse order. None of the four subjects showed any real loss of efficiency in the second period as compared with the first except in the speed of movement test. This was the only test requiring vigorous physical exertion. Nevertheless the feeling of fatigue induced by the ninety minutes of uninterrupted, high-tension work was very marked.

3. The hypothesis is advanced that fatigue is a negative emotional appetite. As such it is to be differentiated from the physiological phenomenon of exhaustion of which it is a concommitant. As an emotion in the broader sense, fatigue is a conscious (if negative) motive to action. Fatigue does not directly cause work-decrement but raises the threshold at which work motives are effective. But if such positive motives are adequate at all, the fatigue—as distinguished from the accompanying exhaustion—has no effect upon work efficiency. This hypothesis deserves experimental investigation since it seems to explain many otherwise puzzling facts.

RELATIVE PERFORMANCE OF COLLEGE STUDENTS AS CONDITIONED BY TIME OF DAY AND DAY OF WEEK

BY DONALD A. LAIRD

Colgate Psychological Laboratory

With the assistance of Ralph K. Hoitsma, B. Appleby, A. J. Powell, Jos. Slosky, W. L. Alcorn and J. Crosby Chapman.

I. Introduction

The alleged rhythmical fluctuations of physiological and psychological processes—seasonal, daily, and hourly—has occupied the attention of some workers in these fields for many years. Altogether a great deal of experimental work has been done in investigations that, when critically considered, are far from decisive. It would be of much practical and theoretical value to know in detail about these fluctuations in performance that appear to be conditioned—remotely, perhaps—by chronological and meteorological factors.

II. Historical

Excellent summaries of previous work are given by Marsh [1] and by Gates.[2, 3] Other reports that have appeared since these summaries and with which the worker interested in this field should be familiar are Huntington's observations recorded in "Civilization and Climate," Baumberger's reports on fluctuations in industrial output, Muscio's report of experiments on women students in the *British Journal of Psychology* in 1920, and the data in the occasional bulletins of the British Industrial Fatigue Research Board.

The previous work may be summarized by stating that there is but little agreement in the findings of those who worked with test materials, while the data gathered under industrial conditions represent fairly homogeneous results.

[1] *Archives of Philos., Psychol. and Sci. Methods*, July, 1906. No. 7.

[2] Univ. of Calif. publications in psychol., 1916, Vol. 2, 1–156.

[3] Univ. of Calif. publications in psychol., 1916, Vol. 1, 323–344.

The lack of agreement in the former may be due to a lack of experimental insight rather than to the absence of any consistent variations in performance from time to time. Many of the conclusions of the experimental workers are forced to say the least and not substantiated by the reliability of their data, nor confirmed by others.

III. Experimental

The present experiments were carried out with 112 college students serving as subjects. These students were divided into seven squads of 16 each. The personnel of each squad was so selected that the distribution of intelligence scores from squad to squad was as nearly as possible the same.

A schedule for experimentation, as detailed in the first table, was so arranged that the performance of all 112 of the subjects could be averaged, keeping practice, hour of the day, and hour of the week constant. Inspection of the schedule will show that when the records for 8 A.M. are averaged one seventh of these are taken on Sunday, one seventh on Monday, and so on through the days of the week. With reference to practice it will also be seen that in the 8 A.M. records one seventh are for the first practice, one seventh for the second practice, and so on.

Later, when the records are taken to contrast the performance on the different days of the week, it will be seen that one seventh are for 8 A.M., one seventh for 10 A.M., and so on for the seven hours sampled. Practice, also, is distributed as when the hourly records are tabulated.

The particular hours selected were chosen to give a comparison with the characteristic points of the curves obtained by Gates and in addition to compare the performance of college students in tasks similar to study problems in the hours of the evening with the hours of the school day.

In order to carry out this schedule a period of six weeks experimentation was necessary. The subjects were 'fogged' as to the real issue of the experiment by being given the impression that it was connected with a study of temperature changes.

In selecting the tests to be used the aim was to use those that appeared on their face to be probing responses similar to those made by the student in preparing for classes. The tests used were:

(*a*) Immediate retention of selections read (Dutch Farmhouse, et cet.).
(*b*) Reading time of (*a*) when read *intentionally* to remember.
(*c*) Two column addition (Woodworth and Wells).
(*d*) Substitution test (Starch).
(*e*) Single column addition (Woodworth and Wells).
(*f*) Part iii of Thorndike Intelligence Test ('Comprehension').
(*g*) Delayed retention of selection read (*a*).

The tests were given in the order listed above as group exercises. The time consumed for each battery was on the average fifty minutes. The hours studied, therefore, are in reality 8:00 to 8:50, 10:00 to 10:50, and so on.

The tests were given in seven different forms, but not equated forms. Since each form is equally distributed as regards practice and time of day and day of week this policy will be reflected in the data only in raising the standard deviations.

During the period of the experiments the students who acted as subjects did not vary from their usual routine, except to report for the tests at the time scheduled for their particular squad. It should be stated that the 112 students whose records are used served without missing a single appointment. A few others started, and indeed saw through, the experiment, but inasmuch as one hour of more had been missed their entire records were discarded. This was necessary since the schedule for practice and the other controlled variables would have been seriously impaired if they had been included.

TABLE I

SCHEDULE OF THE TEST PERIODS OF THE SQUADS

Squad A

Day	Week[1]	Hour
Saturday	2	10 P.M.
Thursday	3	9 P.M.
Tuesday	3	8 P.M.
Sunday	4	4 P.M.
Friday	5	1 P.M.
Wednesday	5	10 A.M.
Monday	6	8 A.M.

Squad B

Day	Week	Hour
Sunday	1	8 A.M.
Friday	3	10 P.M.
Wednesday	3	9 P.M.
Monday	4	8 P.M.
Saturday	5	4 P.M.
Thursday	6	1 P.M.
Tuesday	6	10 A.M.

Squad C

Day	Week	Hour
Monday	1	10 A.M.
Saturday	2	8 A.M.
Thursday	4	10 P.M.
Tuesday	4	9 P.M.
Sunday	5	8 P.M.
Friday	6	4 P.M.
Wednesday	6	1 P.M.

Squad D

Day	Week	Hour
Thursday	1	4 P.M.
Tuesday	1	1 P.M.
Sunday	2	10 A.M.
Friday	3	8 A.M.
Wednesday	4	10 P.M.
Monday	5	9 P.M.
Saturday	6	8 P.M.

Squad E

Day	Week	Hour
Friday	1	8 P.M.
Wednesday	1	4 P.M.
Monday	2	1 P.M.
Saturday	3	10 A.M.
Thursday	4	8 A.M.
Tuesday	5	10 P.M.
Sunday	6	9 P.M.

Squad F

Day	Week	Hour
Saturday	1	9 P.M.
Thursday	2	8 P.M.
Tuesday	2	4 P.M.
Sunday	3	1 P.M.
Friday	4	10 A.M.
Wednesday	4	8 A.M.
Monday	6	10 P.M.

Squad G

Day	Week	Hour
Sunday	1	10 P.M.
Friday	2	9 P.M.
Wednesday	2	8 P.M.
Monday	3	4 P.M.
Saturday	4	1 P.M.
Thursday	5	10 A.M.
Tuesday	5	8 A.M.

[1] Week of the experiment.

IV. DATA AND DISCUSSION

(*a*) *Presentation of Data by Days of the Week*

In Tables II. and III. are given the data relative to the performance of the subjects in the tests used, with practice

and time of the day equally distributed for each day. The first Table gives the mean score of the 112 subjects together with the standard deviation of the mean. In Table III. are given the differences in performance from day to day in terms of the means, together with the standard deviation of this difference, using the Sunday performance as the basis for comparison. The data of the latter table can be used in critically evaluating the findings of this section.

TABLE II

SUMMARY OF DATA FOR DAY OF THE WEEK

	1[1]	2	3	4	5	6	7	8	9
Sun.....	M 53.1	45.6	7.3	43.4	58.9	25.6	24.8	45.5	43.7
	o 16.9	8.7	3.9	13.1	19.6	10.1	9.4	21.0	19.7
Mon.....	M 56.6	44.0	9.0	53.4	51.7	25.3	24.3	52.0	55.0
	o 14.2	12.2	3.9	20.0	15.3	9.9	10.0	19.8	17.0
Tue.....	M 65.4	46.7	8.1	51.2	48.2	25.7	26.4	45.5	43.9
	o 22.7	10.6	3.7	22.3	15.4	8.6	9.1	17.4	15.5
Wed.....	M 56.1	43.5	7.7	56.8	46.6	33.6	27.6	41.6	41.0
	o 14.2	15.4	3.9	29.1	13.6	14.1	12.4	14.3	18.2
Thur.....	M 49.7	43.7	8.2	45.2	47.7	27.1	26.8	50.6	55.0
	o 14.5	8.3	5.5	14.1	13.2	10.6	11.1	15.1	17.3
Fri.....	M 55.5	44.4	7.0	49.5	51.4	28.6	28.3	53.5	54.5
	o 13.1	11.5	3.4	23.7	12.8	12.3	11.7	18.8	17.2
Sat.....	M 50.7	42.4	7.7	42.7	52.7	26.2	26.5	50.1	48.8
	o 13.9	11.7	4.3	17.5	13.0	10.0	11.1	17.8	19.2

[1] These numerals refer to the tests as follows:

1. Simple addition. Mean given represents number added in two minutes.
2. Difficult addition. Mean represents number added in four minutes.
3. 'Comprehension.' Mean represents score according to Thorndike's method.
4. Substitution. Mean represents letters correctly translated.
5. Reading time. Mean represents seconds for reading logical memory selection.
6. Immediate memory. Mean represents 'ideas' of selection reproduced.
7. Delayed memory. Mean represents 'ideas' reproduced after forty minutes interval.
8. Immediate memory. Mean represents irrelevant words used in reproduction.
9. Delayed memory. Mean represents irrelevant words used in reproduction after forty minutes.

Charts 1 to 4, inclusive, show the varying performance for each of the tests and for the combined tests. Explanation is needed of the way in which the data were treated in ob-

taining the material for the charts. In order to have each test in units that would enable comparison with other tests and also make possible the summing of these scores they were reduced to a common basis by the following formula:

$$\frac{700}{\Sigma \text{ scores}} = k,$$

in which the scores of the nine tests for any given day were summed and a constant determined by which each individual

TABLE III

INCREASE OR DECREASE (−) OVER SUNDAY AS A BASE FOR EACH OF THE TESTS, TOGETHER WITH THE STANDARD DEVIATION OF THE DIFFERENCE

	1[1]	2	3	4	5	6	7	8	9
Mon. Diff.	3.5	−1.6	1.7	10.0	7.2	−0.3	−0.5	−6.5	−6.3
σdiff.	2.1	1.4	0.5	2.2	2.3	1.3	1.3	2.7	2.4
Tue. Diff.	12.3	2.1	0.8	7.8	10.7	0.1	1.6	0.0	−0.2
σdiff.	2.6	1.3	0.4	2.4	2.3	1.2	1.2	2.6	2.4
Wed. Diff.	3.0	−2.1	0.4	13.4	12.3	8.0	2.8	3.9	2.7
σdiff.	2.1	1.6	0.5	3.0	2.2	1.6	1.4	2.4	2.5
Thu. Diff.	−3.4	−1.9	0.9	1.8	11.2	1.5	2.0	−5.1	−11.3
σdiff.	2.1	1.1	0.6	1.8	2.2	1.4	1.4	2.4	2.5
Fri. Diff.	2.4	−1.2	−0.3	6.1	7.5	3.0	3.5	−8.0	−10.8
σdiff.	2.0	1.3	0.4	2.5	2.2	1.5	1.4	2.7	2.5
Sat. Diff.	−2.4	−3.2	0.4	−0.7	6.2	0.6	0.7	−4.6	−5.1
σdiff.	2.0	1.4	0.5	2.1	2.2	1.3	1.4	2.6	2.6

[1] See footnote to Table II.

test score could later be multiplied to give a total score of 700 for the course of the week. A test which by this procedure gave an adjusted score of 100 would be at the level of the average of the week, one below 100 would be below the average of the week, while one in excess of this figure would indicate superior performance. The use of a procedure such as this makes possible the later summing of the tests on an equitable basis. The average performance of the week is also taken as the standard by such procedure rather than any single and arbitrarily chosen day. The horizontal divisions of charts 1 to 4 represent the adjusted scores or 'efficiency'

in comparison with the average for the week, which is indicated by the heavier line at 100.

It should also be noted that in the records of reading time and irrelevant words in the retention tests a large score represents decreased performance. Accordingly in the charts and in Table III. the measures of these tests have been reversed to indicate their true significance.

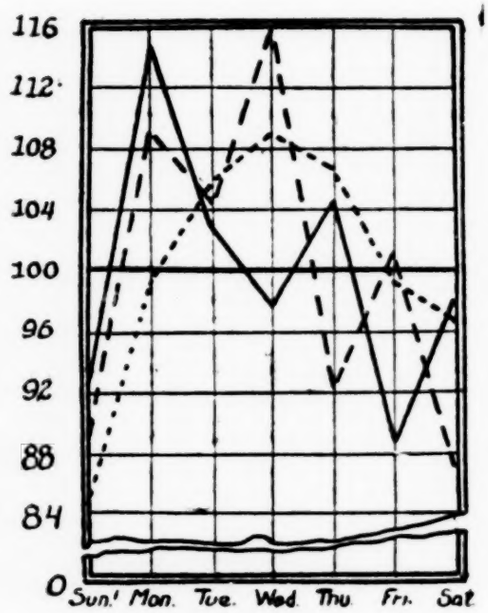


CHART 1

Solid—Comprehension
Broken—Code
Dotted—Reading

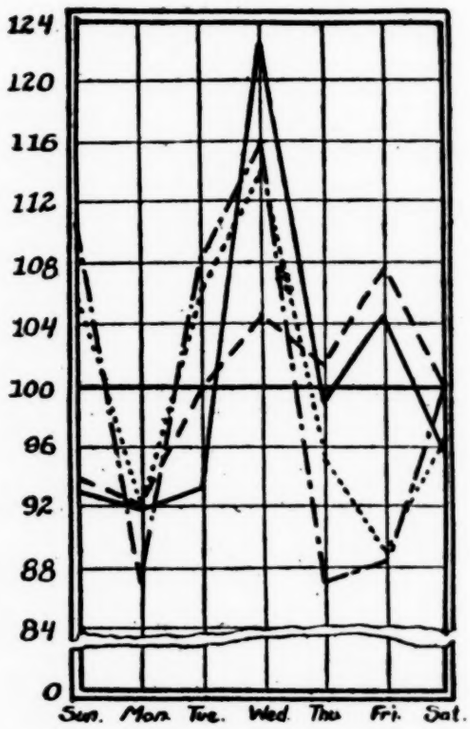


CHART 2

Solid—Immediate Memory, Ideas
Dotted—Immediate Memory, Words
Broken—Delayed Memory, Ideas
Dot Dash—Delayed Memory, Words

Examination of the performance curves of the individual tests shows a general agreement in a rise toward the middle of the week. The peak for the performance in the 'comprehension' test is on Monday; in comparison with Sunday's performance this gain is three times the standard deviation of the difference. (See Table III.) The difference in the absolute score is a gain of 1.7, but in terms of the average day's performance the difference between Sunday and Monday in this particular task is 20 per cent.

The high point in the performance in the addition tests is found on Tuesday—a day later than for 'comprehension.' The daily variation in the two column addition is not significant when the standard deviations are taken into account. The high point of the curve for single column addition, however, is significant. In terms of additions it means that on Tuesday the average performance of the 112 subjects studied would have enabled them to do 369 more additions

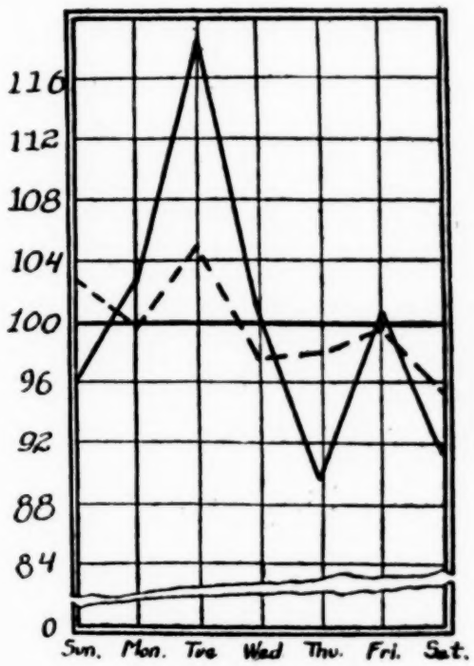


CHART 3

Solid—Single column addition
Broken—Two column addition

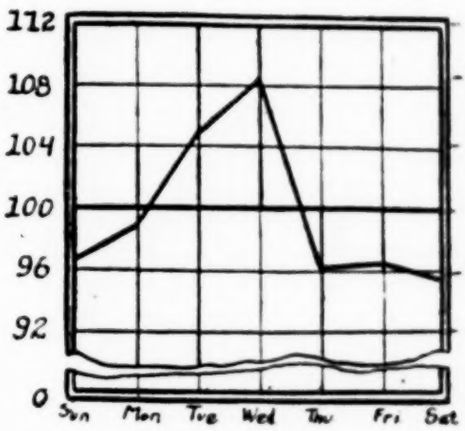


CHART 4.

Results on all tests showing the performance at different days of the week

per hour than on Sunday, assuming that the rate of the first few minutes is continued without change for the longer period.

The peak of the curve of the code or substitution test, of the retention tests, and of reading time occurs on Wednesday. The increase over the average daily performance in each of these is significant when the standard deviation is considered except in the case of the delayed retention of the important points of the passages read. In this later test the peak is found on Friday.

The two troughs in the curves for irrelevant words intro-

duced in writing the memory selection—Monday, and Thursday and Friday—may be taken as indicative of excessive verbiage at these times. The first trough is accompanied by a dearth of points remembered, but this does not hold for the trough later in the week.

Inspection of the first three charts gives the impression of the performance being at its highest around the middle of the week. When the adjusted scores of the individual tests for any day are averaged we obtain the composite picture of daily variance in test performance that is shown in Chart 4.

It is possible to obtain the standard deviation of each of the points represented in this chart by the use of Yule's formula for obtaining the standard deviation of a sum, given on page 211. To obtain this figure it is necessary to know the correlation between the tests. Drawing on the literature of tests for an approximate correlation figure, rather than calculating it from the data, we may be safe in assuming a correlation of .50 to be high. Objections to this correlation being too high may be tempered by reflection that a high correlation penalizes in this particular formula. Proceeding with this assumed correlation figure we find that the standard deviation of each of these points may be taken as not greater than $\pm$ 2.0. Evaluating the data of Chart 4 with this figure we find that the Wednesday high point is four times the standard deviation of the difference between Wednesday and the weekly average, and six times the difference between Wednesday and Saturday, the lowest day of the week in the performances studied.

Considered in the light of the error of sampling it may be said that Chart 4 gives a fair picture of the weekly course of performance in the tests used, namely, a gradual increase in performance during the first four days of the week with an abrupt decline on Thursday that is not altered appreciably until Tuesday. Thursday, Friday, Saturday, and Sunday are the low days for the students studied as a group.

This finding is in general agreement with the data obtained from industrial records. In these latter records, however, it should be mentioned that there are occasional individual

records in which a superior workman increases steadily in output throughout the week. Thus in place of the decrease after Wednesday these few workers studied increase steadily in output throughout the week. It may be that there are individual variations of a similar nature that are masked in our group results. The present data, however, are not such that it is possible to check this very probable assumption. This is an important phase to be investigated, but it is impossible to get at it satisfactorily with volunteer subjects.

It is also possible that if a single work test were used for the whole period of performance different results might be obtained, although the industrial data would give but little indication of such a tendency. This cannot be answered definitely, however, until an experiment similar to the one reported is undertaken with the continuous exercise of a single problem rather than a period in which nine different tests are used. It may very reasonably be argued that the novelty due to the shift from one test to another may offset decreased potential performance.

(*b*) *Presentation of Data by Hours of the Day*

The following tables and charts present the findings for the variation in performance by hour of the day, with day of the week and practice kept constant from hour to hour.

Examination of Charts 5 to 7 will show that there is less agreement from test to test than was found in the data for daily variations. In practically all of the tests, however, 10 A.M. is below 8 A.M., which is at variance with Gates' findings on the memory fluctuations of college students.

When the composite scores are obtained, as shown in Chart 8, a little more resemblance to order is noticeable. There is a steady decline in performance until 4 P.M. This decline amounts to about 10 per cent. less than the performance at 8 A.M., which is five times the standard deviation of the difference that is being used. Following the evening meal there is a slight increase in the performance, which is lost again, however, in the period from 10 until 11 P.M.

It may be that there are also individual differences in

these data that are masked in our search for the variation in the case of the hypothetical 'typical individual.' It may

TABLE IV

SUMMARY OF DATA FOR HOUR OF THE DAY

	1[1]	2	3	4	5	6	7	8	9
8 A.M. M.....	56.6	43.7	8.8	55.1	53.1	31.7	29.5	47.6	49.9
o	14.7	12.5	3.9	27.3	13.9	9.5	9.2	11.1	15.8
10 A.M. M.....	56.1	44.7	8.7	50.8	52.6	27.4	30.0	52.2	53.6
o	15.1	11.4	3.9	27.5	15.5	11.6	10.6	21.4	23.3
1 P.M. M.....	60.7	47.1	6.4	51.2	48.4	28.3	30.6	51.6	51.6
o	16.2	10.8	3.8	22.7	14.0	10.0	12.6	17.9	14.8
4 P.M. M.....	57.4	43.1	7.3	46.2	54.0	25.0	23.8	49.6	51.8
o	25.4	12.2	4.6	18.4	15.6	10.1	9.5	18.3	16.4
8 P.M. M.....	55.5	45.5	7.7	49.7	48.5	22.6	20.9	43.1	46.2
o	13.5	12.3	4.0	19.6	13.9	10.5	10.1	15.9	18.2
9 P.M. M.....	50.9	43.8	9.0	47.3	47.6	27.3	27.2	52.2	51.7
o	11.8	10.8	4.5	22.9	15.9	8.2	17.8	21.1	18.4
10 P.M. M.....	50.8	45.1	7.0	42.9	54.2	25.5	25.5	45.5	43.1
o	12.6	10.3	4.5	16.1	17.1	11.2	10.0	20.3	17.3

[1] See explanation in note to Table II.

TABLE V

INCREASE OR DECREASE (−) OVER 8 A.M. AS A BASE OF THE TESTS, TOGETHER WITH THE STANDARD DEVIATION OF THE DIFFERENCE

	1[1]	2	3	4	5	6	7	8	9
10 A.M. Diff.....	−0.5	1.0	−0.1	−4.3	0.5	−4.3	0.5	−4.6	−3.7
o_{diff}	1.9	1.6	0.5	3.7	1.9	1.4	1.3	2.3	2.7
1 P.M. Diff.....	4.1	3.4	−2.4	−3.9	4.7	−3.4	1.1	−4.0	−1.7
o_{diff}	2.1	1.5	0.5	3.4	1.9	1.3	1.5	2.0	2.0
4 P.M. Diff.....	0.8	−0.6	−1.5	−8.9	−0.9	−6.7	−5.7	−2.0	−1.9
o_{diff}	2.8	1.6	0.6	3.1	1.9	1.3	1.2	2.0	2.2
8 P.M. Diff.....	−1.1	2.2	−1.1	−5.4	4.6	−9.1	−8.6	4.5	2.7
o_{diff}	1.9	1.6	0.5	3.2	1.8	1.3	1.3	1.9	2.3
9 P.M. Diff.....	−5.7	0.1	0.2	−7.8	5.5	−4.4	−2.3	−4.6	−1.8
o_{diff}	1.8	1.5	0.5	3.4	1.9	1.2	2.8	2.2	2.3
10 P.M. Diff.....	−5.8	1.4	−1.8	−13.2	−1.1	−6.2	−4.0	2.1	6.8
o_{diff}	1.8	1.5	0.5	3.0	2.1	1.4	1.3	2.2	2.2

[1] See footnote to Table II.

also be, again, that an hour spent in a uniform task would give results different from an hour spent in working at several

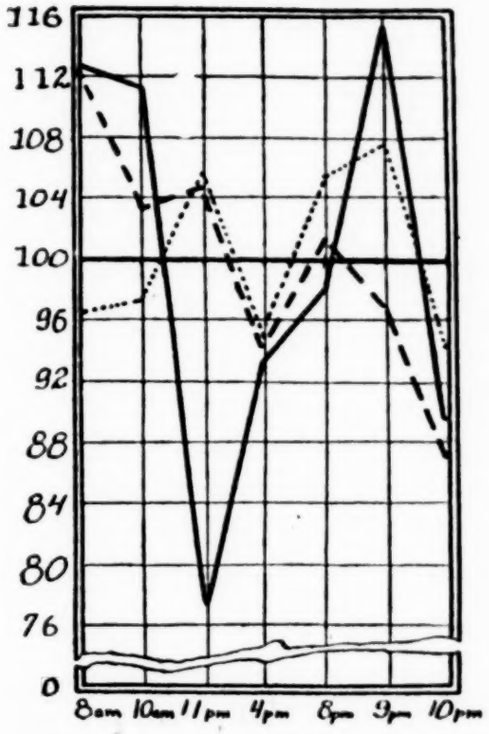


CHART 5. Same plotting as Chart 1

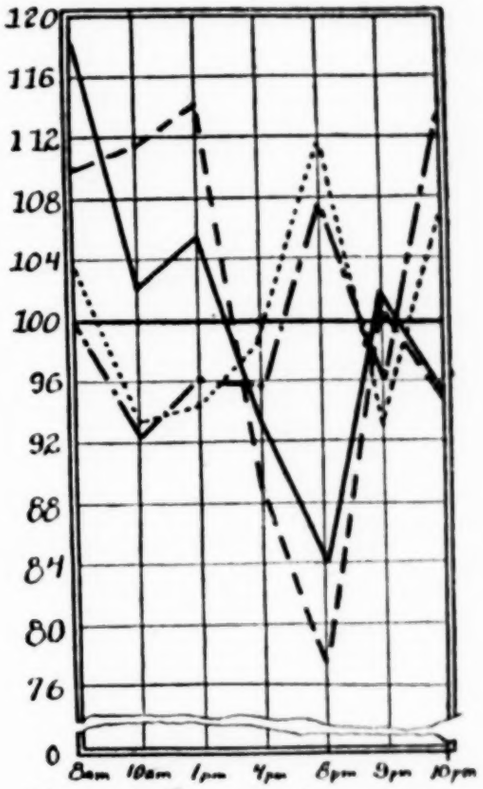


CHART 6. Same plotting as Chart 2

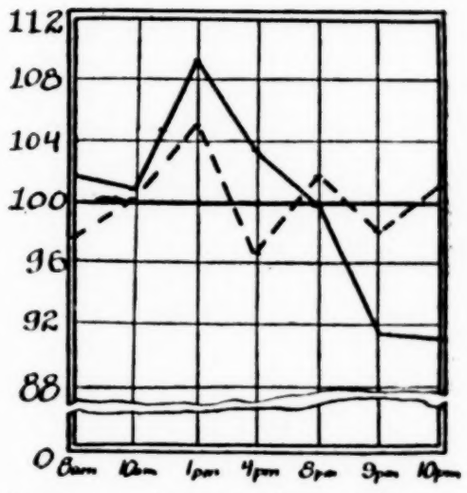


CHART 7. Same plotting as Chart 3

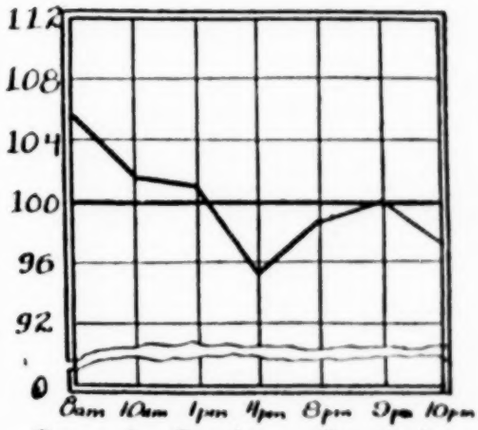


CHART 8. Results on all tests showing the performance at different hours of the day

different tasks. This is indicated in the present data by the lack of uniformity in the hourly curves from test to test.[1] This same absence of uniformity from test to test may give

[1] Data on this are now being gathered in the Colgate laboratory by Frank Cooper.

some insight into possible causes back of the disagreement of earlier investigators. If we had limited the tests to the ones of addition we would have said that the performance was about 7 per cent. above the average for the day at 1 P.M., whereas the total series of tests shows that in reality performance is about 5 per cent. below the average for the day at that hour.

V. Summary

Tests of speed of reading for understanding, grasping of important points from selections read, retention of important items from reading after a few minutes and for forty minutes, were given to 112 college students according to a schedule that made it possible to compare their performance on these tests either for the days of the week or for the hours of the day.

It has been found that the performance was at its peak on Wednesday, that it gradually approached the high point on this day and abruptly declined on the day following to a very low level. In general all of the individual tests followed this curve of rise and fall.

From 8 A.M. until the period from 4 to 5 P.M. there was found a steady decline at the hours studied. From 8 P.M. until 10 P.M. there is a rise, which, however, scarcely comes up to the average of the day. From 10 to 11 P.M. there is a loss.

These conclusions apply, not necessarily to every individual student, but represent accurately the average performance of a large group, or the 'typical student.' It may perhaps be that there are individual exceptions to these findings as the result of habits of study and work. It is doubtful if there are variations as a result of one student being intrinsically and apart from habits of work different from others.

The low correlations between tests and apparent unreliability of tests in other investigations may possibly be explained largely by the present findings. An addition test on Tuesday will be found to vary as much as 25 per cent. above one given on Thursday due to performance being in a 'higher gear' on the former day. Similarly with respect to

hours of the day an addition test from 1 until 2 P.M. may be found to give results 7 per cent. better than the same test at 8 to 9 A.M. We have found that not all tests vary in the same direction or amount, so the general curves we have presented cannot be used to correct for these variations. It is important, however, that this influential source of unreliability in mental test work be given its due emphasis.

It should be given further emphasis that this variability in performance does not necessarily represent changes in efficiency. Although the phrase mental efficiency is deftly bantered around no one has, strictly speaking, yet made a study of mental efficiency. What has been studied in varying output under varying conditions. The input is still to be measured, and until that is known for each of the different performances no one can say which is the more efficient time or mode. It may very well be that conditions for maximum performance are in fact the least efficient when energy consumption is taken into consideration. The past year, under the auspices of the National Research Council, I gathered data on this point which indicates in its present form that there may be marked loss in 'mental efficiency' although the performance remains the same or may actually increase.

Out of statistical curiosity it may be stated that in reaching the findings reported in this paper six weeks of time was needed to collect the data. There were 4,704 test blanks to be mimeographed and scored. In addition to this there were around 17,000 numbers to be added, 26,000 numbers to be squared, and 53,000 subtractions to be made. All the computations had to be checked.

REACTIONS TO VISUAL STIMULI IN AFFECTIVE SETTINGS

BY F. L. WELLS

Boston Psychopathic Hospital

These experiments were undertaken with a view to learning something about the latent period of affective reactions. The question at issue was, how long after the exhibition of a stimulus calculated to arouse some affective reaction, does it take for an affect to be distinguished. The problem is like that undertaken by Nakashima, but the method differs in at least one important particular. Nakashima's subjects indicated by a conventional reaction, when the affect was observed, regardless of its character. In the present experiments, the subject is asked to distinguish whether the affect is pleasant or unpleasant, by two different reaction movements, in most cases operation of telegraph keys by right and left hands.

The present report is concerned with visual stimuli (pictures), printed or mounted on cards about $3\frac{1}{2}$ by $5\frac{3}{8}$, which could be conveniently shown in a Dodge mirror tachistoscope, operating as an exposure screen. As in previous studies, a ballistic galvanometer served as a chronoscope. The generally used stimuli were women's faces, with the Binet pictures as a starting point, but other types of picture were used as subsequently described. Visual stimuli are advantageous in this inquiry, in that the incidence of stimulus and reaction are relatively easy to fix and time. They are disadvantageous in that the affects are generally feeble and conventional. Indeed, though the setting was always affective, and it was always a question of relatively liking or disliking a picture, it soon developed introspectively that the reaction might be determined by factors that it would be straining a point to call affective. It would appear that the subject had designated a stimulus agreeable simply through recognizing the outlines shown as something conventionally

regarded as agreeable, but without having actually felt it as such. This was a disadvantage to which these visual stimuli were especially subject. After it developed, it was endeavored to meet it by taking introspective notes on each reaction, as to whether the reaction seemed to have been essentially determined by feeling, or whether it was more a matter of judgment. Naturally, these distinctions were not always clear-cut; but enough reactions are definitely assigned to the affective group to permit of some generalizations on the topic in hand.

For such purposes as these, only persons accustomed to some observation of their mental processes are serviceable as subjects. Most of the observers were graduate students in psychology, or of an equivalent order of competence.

In the first experiments two faces, such as the Binet pairs, were exposed on a single card; and the observer was requested to raise the telegraph key on the side where the 'prettier' face appeared. This was not very satisfactory, as disturbing eye movements were observed with the faces as close together as practicable. The method was then tried of showing single pictures, either 'pretty' or otherwise with degrees of difference of about the same order as those in the Binet pictures. The subject would then raise the right hand if the face was good-looking, the left if it was otherwise. The disadvantage of this procedure was that it was inconvenient to secure a considerable array of sufficiently contrasting portraits. The plan was then adopted of taking from a number of stimuli, one that seemed about midway, and using this as a standard. It was placed in one of the apertures of the Dodge apparatus, so that the observer could see it all the time except when one of the varying stimuli was being exposed for comparison. The instruction was then to raise the right hand if the new stimulus were more agreeable than the standard one, the left if otherwise. This was the most generally serviceable method, adapted also to other types of pictures, as landscapes on picture postcards, and it was the most used.

The conduct and results of these experiments are described with somewhat more detail in the following.

a. Experiments with Pairs of Pictures

As noted, the first observations were made with the Binet pictures. These are here denoted by numbers. From top to bottom as printed on the Stoelting card 78601 (Point Scale), the three better looking pictures are denoted as 1, 3, 5, the three less attractive pictures as 2, 4, 6. Five trials were made on each of two subjects, with each of the combinations 1–2, 3–4, 5–6, 2–4, 2–6, and 4–6. Each single trial alternated with the exposure of a card in which black and white were substituted for the two pictures. This was in order that the reaction times to the pictures might be compared to that with an objective choice whose properties were already better known.

For one of the observers measured, the discrimination time for the black and white stimuli averaged 411 sigma. The average time for recording the decision as to the prettier of the two pictures was 763 sigma. In the other observer these figures were respectively 364 and 839 sigma.

It is generally accepted that with sensory stimuli, the time of perceiving a difference is a reasonable measure of the magnitude of the difference itself. A difference discriminated in a shorter time is a greater difference. The question is at once raised, whether a similar principle applies in the affective sphere. It is not to be supposed that the differences in prettiness between the pairs observed are equal. The differences will, moreover, change in value for different people; some persons are more sensitive to particular kinds of prettiness than others. Does the speed with which one arrives at a decision of this kind measure the magnitude of the affective reaction? If so, persons of major affective sensibility should give relatively short choice times in experiments of this type.

There are also possibilities for measuring the general affective value of different types of stimuli. Thus, both observers find relatively great difference between pictures 5 and 6; that is, their choice reaction times are short. On the other hand their reaction times to pictures 1 and 2 are among the longest for these observers. This would indicate little difference in their objective value, and a presumably slight

affective value for each. In this connection it may be noted that the observer with the shorter choice times to black and white, had the longer choice time to the pictures. The observer with the longer affective judgments, had the shorter sensory ones.

A few observations were made vocally, the observer naming the side on which the 'prettier' picture appeared, or the side on which the black or white appeared. These showed similar relation between the two observers.

Other pairs of pictures were prepared, from reproduction of photographs of celebrated women. The differences were probably not so pronounced as in the Binet pictures, but the faces were brought closer together. The time was markedly reduced, and this, together with subjective disturbances from eye-movement, led to the early abandonment of exposure in pairs.

b. Experiments with Single, Contrasting Pictures

A more elaborate experiment was then carried out with four observers. A series of twenty presumably beautiful and twenty presumably plain women's faces were mounted on suitable cards. There was similarly prepared a series of ten men's faces; these were selected without regard to appearance; all are more or less prominent business men, in ordinary business dress. In these experiments, no face was recognized by a subject as to name. There was also prepared a series of ten adjectives among the simplest for an opposites test. Prior to the experiment, ten of the men's faces were presented to the observer, with their names, and these the observer was requested to learn until all these names and faces could be associated without the least apparent difficulty. Reactions to black and white, and a simple reaction to sound with the right hand, were also employed. There were in all nine types of reaction in the experiment. A single one of each type was made, in a serial order, and after the completion of such a cycle, it was repeated in reverse order. This process continued until ten reactions of each type had been obtained. This was done once only for each subject; in the

nature of the experiment, successive repetitions are not comparable. Verbal instructions as to each reaction were given as necessary, and throughout, as in a previously reported experiment, there hung before the observer a schedule showing the order and nature of the successive reactions. The reader may better understand the detail of the experiment through a reproduction of this schedule:

Item

1. If you *recognize* the face shown, raise the *right* hand.
   If you do *not* recognize it, raise the *left* hand.

2. If you *recognize* the face shown, say '*yes*.'
   If you do *not* recognize it, say '*no*.'

3. If this picture represents a physically *attractive* woman, raise the *right* hand.
   If this picture represents a physically *unattractive* woman, raise the *left* hand.

4. If this picture represents a physically *attractive* woman, say '*yes*.'
   If this picture represents a physically *unattractive* woman, say '*no*.'

5, 6. If a *white* card is shown, raise the *right* hand.
   If a *black* card is shown, raise the *left* hand.

7. Give the *opposite* of the word shown.

8. *Read* the word shown.

9. Sound stimulus, click of relay. Give the word you just gave as a simple reaction.

10. Sound stimulus, click of relay. Release the *right hand* key.

For the present remarks, the center of concern is the material of items 3 and 4; the remainder were introduced for comparison and as controls. The observers in this experimental series were all men, and the experimental task in items 3 and 4 was formulated in terms of the attractiveness of the woman represented, as an erotic object. In items 1 and 2 they were not asked to name the man whose face was shown, but merely to indicate if the face was recognized as seen in the previous examination of the faces. The setting of items 3 and 4 is affective, that of all other items in the test is sensory or cognitive; but items 1 and 2 were intended to parallel otherwise 3 and 4 as closely as practicable.

The central results of these experiments, as expressed in the averages and mean variations of the different observers for series 3 and 4, are shown in the following table: (time units are all sigma)

TABLE I

Series	Observers				
	I	II	III	IV	V
3.....	728	833	656	588	816
m.v.....	91	128	73	76	156
4.....	653	753	755	524	781
m.v.....	84	125	91	48	194

The average times observed range between 500 and 850 sigma, a central score being about 700 sigma. These findings are thus in substantial agreement with Nakashima's, despite the difference in procedure. Subjectively, however, there are represented in these results two types of reactions, those in which the reaction seemed to introspection determined by affective factors, on the other hand by cognitive. No clear difference appears in reaction time according to this distinction, but the matter will for its theoretical interest be taken up in more detail below. The records of only the first four observers named are to be utilized in this respect.

For closer comparisons, the chief results were reduced to a percentage basis. The standard selected for this purpose was the average reaction time to black and white (items 5, 6), the most elementary and objective choice time available. There was computed the per cent. which each of the eight other scores made in the experiment, was of this figure. These results are shown in the following table. (Voice reactions are the better equated with hand reactions by deducting from the voice reaction times that quantity by which the simple voice reaction times are longer than the simple hand reaction times, this also in the previous table.)

The judgments made in the affective setting, series 3, 4, appear shorter than those of the cognitive setting, series 1, 2. When the correction for the voice key has been made, the figures for series 2 and 4 are also shorter than those of series 1 and 3; that is, the discriminations are registered with the voice somewhat more rapidly than with the hand. This has been already commented on; the voice is a more natural medium for rendering such decisions than is the hand. The

shortest times for series 3, 4, are shown by observers III and IV. Observers III and IV are younger men than the others, and to the observation afforded, were persons of more active instinctive life. Such impressions as are called for in series 3, 4, might thus be expected to register more strongly and promptly in them than in the others. A somewhat similar situation appears with regard to series 7. This is the time it takes to give the opposite of a word. The per cent. which this is of the choice time, for black and white, varies from 160 in observer I, to 271 in observer V. The order of these figures corresponds fairly to the relative extent to which the various observers have had occasion to manipulate ideas in verbal terms. This would not be the ranking in an intelligence test; here

TABLE II

Series	Observers				
	I	II	III	IV	V
1.....	153	188	178	172	174
2.....	147	163	118	179	163
3.....	155	159	129	151	155
4.....	139	142	148	135	149
5 and 6.....	100	100	100	100	100
7.....	160	185	190	210	271
8.....	129	116	140	137	122
9.....	(28)	(27)	(36)	(34)	(52)
10.....	28	27	36	34	52

the first place would probably go to observer IV. who is relatively slow in the verbal formulation, but much the fastest in registering the distinction of black and white. With regard to simple reaction time the situation is similar. I. and II. are close together, so are III. and IV., while V. is at the further extreme. With I. and II. the simple reaction is the easier, with III. and IV. the elementary, objective choice reaction is the easier. With observer V. the choice reaction is only twice the length of the simple, which is extremely long. It is noteworthy that the time of reading a word seems unrelated to that of giving its opposite, and not reflecting any special differences in the personalities concerned.

The above points are specially noted because, with the

multiplication of mass psychometric methods involving relatively simple techniques, the rôle of refined laboratory procedure in analyzing the mental life has been thrust into the background. Years ago, Cattell called attention to some very plain indications in this field, which were partially rediscovered and given a considerable though one-sided elaboration, by the psychoanalytic users of the association test. The apparatus which accurate work of this nature demands, places on it limitations which do not exist in the less specialized procedures of psychometrics. On the other hand, the work of Seashore in music, and of Henmon in aviation, are examples of penetration by laboratory methods into fields inaccessible to ordinary psychometric procedures. Some of the present results also, seem related to important factors in the personality as yet outside the scope of 'tests.'

c. Experiments by Comparison with Standards

By the foregoing experiments the general relations of the affectively set reaction to other types of reaction task, seemed to be established. In further experiments accordingly the bulk of the observations was concerned with the affective task and the others not introduced, or only in a minor degree as controls. Practically all these observations utilized the last method to be developed, namely the setting of a standard, with which the exposed picture was to be compared. Reaction was with the right hand if more pleasant, with the left if less pleasant, than the standard. Both men and women participated as observers. Pictures of both men and women were utilized as stimuli. With the women observers, a few observations were also made in which plates illustrating various styles of hats served as stimuli. Some use was also made of picture post cards illustrating landscapes, and these were a very satisfactory type of stimulation for the purpose.

The following table indicates the different observers participating; the type of affective stimulus employed; the number of observations of the type; the average times in sigma of this number of observations; the m.v. of this average. For observations of different types which were part of a single

experimental session, the designation of the observer is not repeated. Observers VI., VII. and VIII. are men; IX. and X. are women.

TABLE III

Observer	Stimuli	No. Obs.	Av. Time	M.V.
VI.	Women's faces	16	1,120	219
VI.	Women's faces	8	773	161
	Men's faces	9	960	143
	Landscapes	10	1,023	95
VII.	Women's faces	17	2,123	496
VII.	Women's faces	10	1,541	538
	Men's faces	10	1,842	641
	Landscapes	10	1,265	275
VIII.	Women's faces	10	1,522	384
	Men's faces	10	1,552	395
	Landscapes	10	1,556	456
IX.	Hats	13	768	131
	Women's faces	16	880	222
IX.	Women's faces	10	697	122
	Men's faces	10	735	125
	Landscapes	10	738	125
X.	Hats	14	965	191
X.	Women's faces	10	750	123
	Men's faces	10	900	180
	Landscapes	10	871	167

The five groups of observations noted in this table, which include women's faces, men's faces, and landscapes, include also observations made with black and white, and simple reactions with right and left hands. The scores in all these observations are summed up on a percentile basis, as was done in Table II. The per cents. which the various types of reaction are of the reaction times to black and white, are as follows:

TABLE IV

Stimulus	VI	VII	VIII	IX	X
Women's faces	188	364	308	146	167
Men's faces	234	436	314	154	210
Landscapes	250	300	315	155	193
Black and White	100	100	100	100	100
Simple, Rt. hd.	39	55	40	49	42
Simple, Lft. hd.	41	66	38	43	32

The observations with women's faces in this table are not

directly comparable to those in Table II., because there the decision was whether the face was attractive or otherwise, and there was intended to be no room for doubt. Here decision is whether the stimulus looks more or less attractive than the standard, and the procedure sometimes reduces the differences to a minimum. The ratios in this table are thus generally greater, often much greater than those in the other. There were three observations in which no decision at all was reached in a period that could be timed. It will be noted that the men's faces always take longer than the women's faces, and that the place of the landscapes is irregular.

In this respect the two women observers show no characteristic differences from the men. The women observers, however, show shorter affective reactions than any of the men, and very much shorter than two of the men. This falls into line with the greater affective sensibility recognized among women, which makes the affective differences in the pictures greater for them. The behavior of the affective reaction time in these observations has been such as strongly to suggest its use as a gauge of affective sensibility in the various fields denoted by the different classes of stimuli. The relatively elaborate technique of these experiments is hardly suitable to the purposes of psychometrics. It is not difficult, however, to devise a chain reaction experiment of the sorting type, from which analogous results might be expected.

The affective process is essentially subjective, and the ultimate criterion of whether it has governed a reaction is the introspective account thereof, by a person whose introspection is accepted. Since it is always possible to question the introspection of any person, there are limits to the objectivity which inquiries in this field can attain. The rôle of suggestion in these introspections is potentially great, and the ultimate bearings of the inquiry were not discussed with the subjects beforehand. The observers probably represent as competent a group as could be got together without preconceived notions that might definitely affect the experimental findings. As already noted, the uniform practise was to take with each

observation a brief introspective note as to whether the reaction had been rather affectively or cognitively determined. The experimenter jotted these down verbatim in shorthand. In the cases where the writer himself served as subject, and on a few other occasions, the introspective notes were written by the subjects themselves.

There are in all 277 reactions of affective setting, for which separate introspective classifications are available. Of these, 179 are classified as predominantly affective, 67 as predominantly cognitive, 31 as doubtful. There were 177 reactions to women's faces, and 50 each to men's faces and to landscapes. The women's faces and the men's faces each yielded slightly over twice as many affective as cognitive reactions. The landscapes yielded 40 affective reactions as against but 6 cognitive and 4 doubtful. In respect to reaction time, no consistent difference appears between the two introspective classifications, though in the majority of cases the cognitive reactions average longer. This is in line with a series of 14 reactions to hats with Subject X., not yet mentioned, on which an introspective note states the reactions as a whole are 'mainly affective, made up mind before had time for (cognitive) judgment.' In various single introspections, the affective reaction is also stated to precede a cognitive process.

As noted, a leading point of inquiry in these experiments, was the lower limit of time within which affective reactions might be found to take place. Evidence on this point is perhaps most clearly presented by quoting the notes of a number of individual reactions, affective to introspection, in connection with the time they took. A few introspections of other types are included for comparison. Such presentation would doubtless gain in effectiveness if the actual stimuli could be reproduced, but this does not seem practicable. Statement of general character is all that can be offered. Items 1–19 below are portraits of women. In the remaining items the nature of the stimulus is stated for each case.

(The times are quoted in sigma, but only to the nearest 10 sigma unit, as a greater degree of precision is under these conditions an artefact.)

Item	Observer	Time	Notes
1.....	I.	580	Purely affective to introspection—a good reaction.
2.....	I.	610	Reaction occurred without conscious concomitants.
3.....	I.	660	Affective elements outweigh cognitive. Certain curves in face give unfavorable affect.
4.....	I.	700	Picture recognized from previous exhibition. Cognitive, no affective elements perceived.
5.....	I.	760	Picture remembered, pure cognitive response to memory of last time seen.
6.....	II.	1,090	Affective; suggested unpleasant disposition, trying to look sweet.
7.....	II.	790	Affective; positive, passing at once into distrust.
8.....	II.	830	Affective, mildly; amused at big mouth.
9.....	II.	520	Affective; excellent for 'soul-kisses.'
10.....	IV.	730	Fairly strong cognitive overbalances faint affective process; mildly positive.
11.....	IV.	520	Affective; rather strong, negative. Feeling came almost instantly. Later cognitive processes somewhat decreased aversion.
12.....	IV.	670	Affective, quite strong; at first. The cognitive elements also strong. Both negative.
13.....	IV.	460	Almost entirely affective, very strong. Little cognitive.
14.....	IV.	500	Strong negative affect which controlled the reaction. The aversion much relieved by cognitive factors.
15.....	IV.	640	Strong affective aversion.
16.....	VI.	1,550	No affect; just from a judgment; a little more pleasant than the other.
17.....	VI.	900	Entirely affective; good impression; that is all there is to it.
18.....	VI.	890	Judgment; some affect about the religious factor; thought of as over-religious.
19.....	VI.	700	Ah!—Purely affective.
20.....	VI.	910	(Man's face.) I react perhaps not so much against the man as against the photograph—purely affective.
21.....	VI.	980	(Landscape.) O lovely—purely affective.
22.....	VI.	960	(Landscape.) Say, isn't that a stunner—purely affective.
23.....	VII.	1,550	(Portrait of woman.) Mostly one of feeling; instantaneous emotional response.
24.....	VII.	1,040	(Portrait of woman.) O my—at first purely affective.

Item	Observer	Time	Notes
25.....	VII.	980	(Portrait of woman.) That was affective . . . quite quickly modified to cognitive, which was almost as quick, not quite.
26.....	VII.	3,620	(Portrait of woman.) Decidedly cognitive, which probably accounts for the long time of the decision.
27.....	VII.	1,560	(Portrait of woman.) Both cognitive and affective elements. Affective at first of course.
28.....	VII.	1,000	(Portrait of woman.) Did not take long to decide that, it is affective.
29.....	VII.	1,590	(Portrait of woman.) Of course that was affective rather than at all cognitive.
30.....	VII.	2,810	(Man's face.) My first impression, which was affective, was better; finally was cognitive judgment. (Reaction was negative.)
31.....	VII.	960	(Landscape.) Ah! Immediately roused feeling that is decidedly better than the standard.
32.....	VII.	1,000	(Landscape.) Another beautiful—practically all these affective. All these landscapes decidedly affective.
33.....	VII.	1,000	(Landscape.) Decidedly affective; right (hand) of course; a remarkable piece of scenery.
34.....	VIII.	1,090	(Landscape.) Affective; afterwards with cognitive aspect judgment reversed.
35.....	IX.	890	(Woman's face.) Affective; the smile was terrible.
36.....	IX.	670	(Woman's face.) More intellectual (cognitive); looked brighter.
37.....	IX.	1,100	(Man's face.) Terrible mouth, oversarcastic—cognitive, intellectual. (Note the adjectives used for a reaction reported as cognitive. Observer is well trained psychologically.)
38.....	IX.	710	(Man's face.) Affective first, then intellectual; all previous to reaction.
39.....	X.	850	(Hat.) Distinctly affective; did not have to stop to judge it; just knew it was pleasant.
40.....	X.	520	(Woman's face.) Entirely affective; liked the smile; did not think of anything at all.
41.....	X.	1,100	(Landscape.) Had a hard time deciding; uncertain as to intellectual or affective factors.
42.....	X.	560	(Landscape.) Entirely affective; liked it better because I had seen it before.

The general evidence of these findings is that affective reactions, even in the minor intensities here dealt with, are not uncommonly established within periods of the order of four fifths of a second. Any process with a longer latency than this, is contraindicated as the physiological substrate of the affective life.